

個別介入効果を評価する推薦システムに関する研究

1X17C024-3 今福太一
指導教員 後藤正幸

1. 研究背景と目的

近年、EC サイト上で多種多様な商品が扱われるようになり、多様化する各顧客の嗜好に合致した商品を提示する推薦システムの重要性が高まっている。推薦すべきアイテムを選定する手法の1つに、協調フィルタリングがある。この手法は、EC サイト上のユーザの購買履歴を用いて各アイテムの購買確率を推定し、購買の可能性が高いアイテムを推薦するものである。しかし、購買確率が高いアイテムの中には、日用品などの定期的に購買されるアイテムも存在しており、そのようなアイテムは推薦の有無に関わらず購買される可能性が高い。つまり、推薦による購買確率の上昇幅（推薦効果）が小さいため、推薦の必要性が低い。推薦できるアイテム数は限られているため、推薦効果が高いアイテムを特定することが推薦システムの課題の一つとなっている。

推薦効果は、アイテムの推薦を介入とみなすことにより、因果推論の分野における介入効果と捉えることができる。ここで、個別のユーザの介入効果を推定するモデルとして Estimating Individual Treatment Effect[1](以下、Estimating ITE)が知られている。例えば、ある薬のある患者に投与するときの血糖値の下降幅を知りたいとき、このモデルを用いると、仮に薬を投与する場合としない場合の血糖値を推定することができ、この差分が薬の介入効果を表す。この例のように、因果推論は1種類の介入を扱う医療や経済の分野での活用が主であり、複数種類の介入を扱うことを対象としない。一方で、EC サイトにおける介入である推薦するアイテムの候補は、ユーザとアイテムの組み合わせの数だけ存在する。そのため、単純に Estimating ITE を推薦システムに適用した場合、推薦アイテム数に応じた個数のモデルが必要になり、計算量が膨大になるため現実的ではない。

そこで本研究では、単一の Estimating ITE を複数の介入における個別介入効果を推定できるよう拡張し、推薦システムに適用する方法を提案する。具体的には、従来用いられてきたユーザの特徴量に加え、アイテムの特徴量も用いることで、全てのユーザとアイテムのペアに対する推薦の効果を単一のモデルで個別に推定する。これにより、購買確率が高いアイテムではなく、推薦効果が高いアイテムを推薦することが可能になる。最後に、人工データを用いた実験により提案手法の有効性を示す。

2. Estimating Individual Treatment Effect

2.1. モデル概要

Estimating ITE[1] の概要を図1に示す。ここで推薦システムにおける共変量を、ユーザやアイテムの特徴量のうち、購買確率と推薦有無の両方に関連するものとする。この手法は、2つのフェーズから成り立っている。1つ目のフェーズでは、介入を受ける群（介入群）と介入を受けない群（対照群）の共変量の分布の乖離を最小化する特徴画像による共

変量の変換を行う。2つ目のフェーズでは、変換した共変量を用いた潜在目的変数の予測を行い、最終的に個別介入効果を推定する。なお、学習モデルには Neural Network が用いられる。

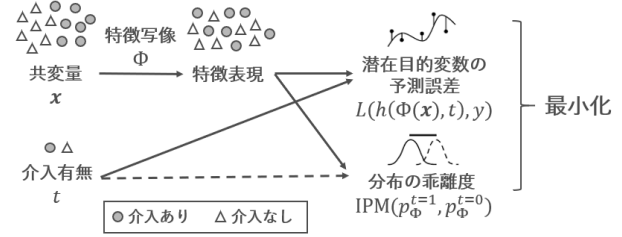


図1: Estimating ITE のモデル概要

2.2. 定式化

ユーザ i の D 次元の共変量を $\mathbf{x}_i \in \mathbb{R}^D$, 介入有無を $t_i \in \{0, 1\}$, 実際に観測される目的変数を y_i^{obs} , 仮に介入を受けないとき ($t_i = 0$), および仮に介入を受けるとき ($t_i = 1$) の潜在目的変数をそれぞれ $y_i^0, y_i^1 \in \mathbb{R}$ とする。このとき、① ($y_i^0, y_i^1 \perp t_i | \mathbf{x}_i$, かつ② $0 < p(t_i = 1 | \mathbf{x}_i) < 1$ であるとする。ここで①は、各ユーザについて共変量 \mathbf{x}_i で条件付けたときに介入有無 t_i と潜在目的変数 y_i^0, y_i^1 が独立になるということ、②は各ユーザが介入を受ける確率が0や1にならないことを表している。また、特徴画像を Φ , 潜在目的変数を予測する関数を h , 介入群と対照群の共変量の特徴表現の分布をそれぞれ、 $p_{\Phi}^{t=1}, p_{\Phi}^{t=0}$, 分布の乖離度を $IPM(p_{\Phi}^{t=1}, p_{\Phi}^{t=0})$ と表す。モデルの学習では、分布の乖離度 $IPM(p_{\Phi}^{t=1}, p_{\Phi}^{t=0})$ と潜在目的変数の予測誤差 $L(h(\Phi(\mathbf{x}_i), t_i), y_i^{obs})$ を最小化する Φ と h を求める。損失関数は式(1)で表される。

$$\min_{h, \Phi} \frac{1}{n} \sum_{i=1}^n w_i \cdot L(h(\Phi(\mathbf{x}_i), t_i), y_i^{obs}) + \lambda \|h\| + \alpha \cdot IPM(p_{\Phi}^{t=1}, p_{\Phi}^{t=0}),$$

$$\text{ただし, } w_i = \frac{t_i}{2u} + \frac{1-t_i}{2(1-u)}, \quad u = \frac{1}{n} \sum_{i=1}^n t_i \quad (1)$$

ここで、 α は分布の乖離度の重み、 λ は正則化項の重みを調整するハイパーパラメータである。

3. 提案モデル

従来 Estimating ITE では、共変量としてユーザの共変量のみを考えるため、介入有無が1つのアイテムの推薦有無に相当し、単一のモデルでは各ユーザに対する1つのアイテムの推薦効果しか推定できない。そのため、膨大なアイテム数を保持するECサイトの推薦モデルとして活用することは難しい。しかし、特徴量はユーザのみではなくアイテムにも存在している。そこで提案モデルでは、ユーザとアイテムのペアの特徴量をモデルの共変量として扱うことにより、1つの介入有無 t で、すべてのユーザとアイテムのペアの推薦有無を表現可能にする。

提案手法のアルゴリズムを以下に示す。

Step1) 共変量の生成

ユーザの特徴量 $\mathbf{x}_u = (x_{u1}, \dots, x_{uD})$ とアイテムの特徴量 $\mathbf{x}_i = (x_{i1}, \dots, x_{iD'})$ を結合し、ユーザとアイテムのペアの共変量 $\mathbf{x}_{ui} = (x_{u1}, \dots, x_{uD}, x_{i1}, \dots, x_{iD'})$ とする。

Step2) モデルの学習

ユーザ数を n 、アイテム数を m とする。観測データ $(\mathbf{x}_{11}, t_{11}, y_{11}^{obs}), \dots, (\mathbf{x}_{nm}, t_{nm}, y_{nm}^{obs})$ について、式 (1) の損失関数の最小化により Φ と h の学習を行う。

Step3) 推薦アイテムの決定

学習したモデルを用いて、各ユーザとアイテムのペアにおける潜在目的変数の予測値 \hat{y}^0, \hat{y}^1 を得る。この潜在目的変数の差分は、各ユーザとアイテムのペアの推薦効果を示す。この差分をもとに各ユーザに対して推薦効果の高いアイテムを推薦する。

4. 人工データを用いた評価実験

提案手法を評価するためには、真の介入効果を正解データとして持つ必要がある。そこで、Causal Tree[2] の人工データの生成過程をもとに購買履歴と推薦履歴を持つ人工データを生成した。これを用いて提案手法の有効性を示す。

4.1. 人工データの生成手順

本研究で用いた人工データの生成手順を以下に示す。

Step1) 特徴量の生成

ユーザとアイテムの特徴量数をともに D とする。また、特徴量はそれぞれ、同じ次元が同じ特徴を表しており、ユーザが好みを自覚している特徴量群 D_1 と好みを自覚していない特徴量群 D_2 が存在すると仮定する。また、EC サイトに存在する人気アイテムを表現するため、推薦されやすいアイテム群におけるアイテムの特徴量のうち、特徴量群 D_1 では、平均値が高くなるように設定する (表 1)。

表 1: アイテムの特徴量 \mathbf{x}_i の生成条件

アイテムの特徴量	D_1	D_2
推薦されやすい群	$\mathbf{x}_i \sim \mathcal{N}(\mathbf{2}_{ D_1 }, I_{ D_1 })$	$\mathbf{x}_i \sim \mathcal{N}(\mathbf{0}_{ D_2 }, I_{ D_2 })$
推薦されにくい群	$\mathbf{x}_i \sim \mathcal{N}(\mathbf{0}_{ D_1 }, I_{ D_1 })$	$\mathbf{x}_i \sim \mathcal{N}(\mathbf{0}_{ D_2 }, I_{ D_2 })$

また、ユーザの特徴量 \mathbf{x}_u は、式 (2) に従って生成する。

$$\mathbf{x}_u \sim \mathcal{N}(\mathbf{0}_D, I_D) \quad (2)$$

Step2) 推薦有無、推薦しないときの購買確率、推薦による購買確率の上昇幅 (推薦効果) の生成

推薦有無、推薦しないときの購買確率、推薦による購買確率の上昇幅はそれぞれ、特徴量群ごとに異なる影響を受けると仮定する。具体的には、推薦有無は特徴量群 D_1 のみから影響を受けるとする。また、推薦しないときの購買確率は特徴量群 D_1 から、推薦効果は特徴量群 D_2 から大きな影響を受けるとする (表 2)。

表 2: 推薦有無等の生成条件

	D_1	D_2
推薦有無	影響 ^有	影響 ^無
推薦しないときの購買確率	影響 ^大	影響 ^小
推薦による購買確率の上昇幅	影響 ^小	影響 ^大

Step3) 潜在目的変数 (購買有無) の生成

推薦しないときの購買確率と推薦による購買確率の上昇幅の和に誤差を足した値を購買確率とする。購買有無は、この購買確率からベルヌーイ分布に従って生成するものとする。

4.2. 人工データの生成条件

ユーザ数を 100、アイテム数を 100、ユーザとアイテムの特徴量数 D は共に 10、推薦されやすいアイテム群は全体の 1 割、特徴量群 D_1 の要素数は 5 とした。

4.3. 実験条件

特徴画像による共変量の変換を行う層と潜在目的変数の予測を行う層はいずれも 3 層とし、各層の次元数は 300 とした。また、潜在目的変数の予測誤差として 2 乗誤差を用いた。比較手法として、ULRMF[3] を用いる。この手法では、まず購買/非購買と推薦/非推薦の観点でアイテムを 4 つのクラスに分け、購買かつ推薦のクラスは「推薦効果あり」、非購買かつ非推薦のクラスは所与の確率で「推薦効果あり」、そのほかのクラスは「推薦効果なし」というようにラベル付けを行う。その後、ラベル付けしたデータを用いて Matrix Factorization を行い、推薦効果のあるアイテムを推定する。

4.4. 実験結果と考察

精度、再現率、F 値の 3 つの指標でモデルの性能を評価した。表 3 に実験結果を示す。

表 3: 人工データにおける実験結果

	精度	再現率	F 値
ULRMF[3]	0.2990	0.1992	0.2391
提案手法	0.3123*	0.5641*	0.4020*

*: $p < 0.05$

表 3 より、いずれの評価指標においても提案手法が従来手法より優れていることが分かる。これは、比較手法が推薦履歴と購買履歴のみを用いているのに対して、提案手法がそれらに加えてユーザとアイテムのペアの特徴量を共変量として用いており、潜在目的変数の予測がより正確に行えているためと考えられる。

5. まとめと今後の課題

本研究では、推薦効果の高いアイテムを推薦するために、個別介入効果を推定するモデルを推薦システムとして活用するための拡張を提案した。そして、人工データを用いた評価実験により提案手法の有効性を示した。今後の課題として、実データを用いた分析などが挙げられる。

参考文献

- [1] Shalit, Uri, Fredrik D. Johansson, and David Sontag, "Estimating individual treatment effect: generalization bounds and algorithms," *International Conference on Machine Learning*. PMLR, 2017.
- [2] Athey, Susan, and Guido Imbens, "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 2016, pp.7353-7360.
- [3] Sato, Masahiro, et al. "Uplift-based evaluation and optimization of recommenders," *Proceedings of the 13th ACM Conference on Recommender Systems*, 2019.