

商品間の代替性を考慮した推薦手法の提案

Proposal of recommender system considering the substitution between items

出雲 英剛*¹
Hidetaka Izumo

佐藤 政寛*¹
Masahiro Sato

園田 隆志*¹
Takashi Sonoda

*¹ 富士ゼロックス株式会社 研究技術開発本部
Research & Technology Group, Fuji Xerox Co., Ltd.

Conventional recommender systems predict purchase probability of an item from estimated user preference of the item. However, non-purchase of an item does not necessarily mean that the user dislikes the item. The user might have bought another substitute and might not need the item anymore. In this paper, we propose an item recommender system that considers the influence of substitution between items.

1. はじめに

ECサイトや店舗の売上向上を目的として、ユーザの嗜好に合わせた商品を推薦する取り組みが行われている。推薦商品を決める手法としては[Sarwar 01]に代表される協調フィルタリングが一般的であり、購買履歴からユーザ毎の推薦商品を抽出する手法として様々な分野で用いられている。また、協調フィルタリングは購買履歴からユーザの商品に対する嗜好を算出し商品の購入確率を予測するため、嗜好を正確に把握することが購入確率の予測精度を向上させる上で重要な要素を占めている。しかしながら、ユーザの未購入商品には嗜好が低いと推測される商品の他に、他の購入商品と代替性を持つ商品が含まれている。[松村 16]では商品の欠品時に発生する代替商品購入行動をモデル化しており、[高橋 92]では牛乳におけるブランド間の代替関係をPOSデータから調査している。このように代替性を持つ商品間においては、一方の購入率が向上した場合に他方の購入率が低下するという特徴があり、嗜好とは無関係に購入率を変動させてしまうため、協調フィルタリングの購入率予測精度を低下させてしまう要因となっている。

本発表では、代替商品の存在を実店舗の購買履歴で検証するとともに自動抽出手法を検討し、代替商品の存在を考慮した推薦手法の提案を行うことを目的とする。

2. 商品購入に影響を及ぼす要因分析

商品購入に影響を及ぼす要因としては、商品の価格や購入商品との代替性、商品への嗜好度が考えられ、ユーザ u における商品 i への嗜好度を R_{ui} 、商品 i の購入率における価格影響度を P_i 、商品 i の購入率が代替商品から受ける影響度を S_i とすると、ユーザ u の商品 i に対する購入意欲度 Pr_{ui} は以下の式のように定式化が可能である。

$$Pr_{ui} = R_{ui} + S_i + P_i \quad (1)$$

ただし、本稿では代替性及び価格の影響をユーザ個人に依存しないものとし、ユーザの商品に対する購入意欲度は購入率 BR_i で置き換え可能であると仮定する。購入率 BR_i は商品 i を購入したユーザの数 b_i 及び b_i は商品 i が購入された日の利用者数である c_i を用いて次のように定義する。

$$BR_i = b_i / c_i \quad (2)$$

しかしながら、購入率 BR_i も代替性と価格の影響を受けるため、購入率を目的変数、代替性、価格影響度を説明変数とした重回帰モデルを作成し分析を行った。代替性と価格影響の偏回帰係数をそれぞれ A 、 B とすると BR_i は次式のように表わされる。

$$BR_i = R_i + A * S_i + B * P_i \quad (3)$$

商品価格の購入率影響度 P_i と商品 i が代替商品から受ける影響度 S_i について、食品スーパーマーケットの購買履歴を使用して検証を行った。使用した購買履歴の概要を以下に示す。

- ・期間:50 日分
- ・商品数:4,073
- ・商品カテゴリ数:12

また、購買履歴の例を表 1 に示す。購買履歴はユーザ、商品コード、商品カテゴリ、商品価格の他に販売日が記録されており、日毎の利用者数を集計する事が可能なデータとなっている。

表 1. 購買履歴のサンプル

ユーザID	商品コード	商品名	商品カテゴリ	商品価格
1	12345	牛乳A	牛乳	200
1	12355	卵A	卵	250
2	12346	牛乳B	牛乳	380
2	12355	卵A	卵	230
2	12365	きゅうりA	野菜	290

2.1 代替商品が商品購入に及ぼす影響の調査

ある商品と同カテゴリ商品の間にはブランド違いや容量違い、産地違い等、特徴が類似した商品が多数存在しており、代替関係にある商品が他カテゴリ内より多いと推測される。今回は、3つのカテゴリ(米類、牛乳、卵)を分析対象として選択した。

選択した商品と同カテゴリ商品間の購入率相関を分析した一例を図 1 に示す。

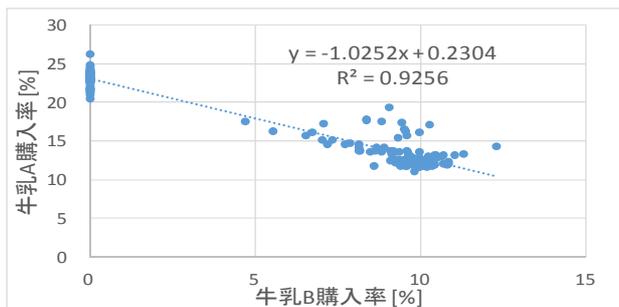


図 1. 代替商品の一例

図 1 では、牛乳 A の購入率が向上した場合に、牛乳 B の購入率が低下するという結果を示しており、同カテゴリ商品内で代替関係が成立している例である。また、この商品の詳細を確認したところ、内容量の異なる商品であった。同一カテゴリ内で代替商品の存在が確認できたため、推薦手法の検討を行った。

3. 提案手法

2 章の分析から、代替商品の購入率向上が他の購入商品の購入率低下を引き起こすことが実データ上でも明らかになったため、この影響を考慮した推薦手法を提案する。

3.1 代替商品の存在を考慮した商品購入率予測

代替商品の存在を考慮した推薦システムのフローチャートを図 2 に示す。

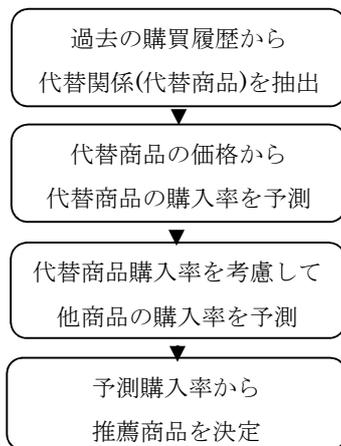


図 2. 代替商品の存在を考慮した推薦システムのフローチャート

本稿では、実験対象商品を牛乳 A と牛乳 B のペアに限定し実験を行った。また、牛乳 B の購入率は購買履歴より算出されているものとする。

3.2 商品購入率予測手法

代替商品が存在した場合に他商品の購入率を予測する手法として、(3)式を利用し牛乳 A の購入率を目的変数、牛乳 A の価格と牛乳 B の購入率を説明変数とした重回帰分析を行っている。

4. 購入率予測実験

4.1 実験概要

分析対象商品牛乳 A について各購入日の購入率を牛乳 A 価格と牛乳 B 購入率から予測する。代替商品は容量違いや味違いの商品が多く含まれていると推測される同一商品カテゴリに限定し予測モデルを作成した。また、牛乳 A 価格のみから予測した購入率予測結果を比較対象とした。

4.2 評価方法

購入率予測の評価方法として、RMSE を用い、牛乳 A が購買されている日の購入率予測を Leave-One-Out で交差検証を行った。

5. 実験結果と考察

購入率予測の実験結果を以下の図 3 に示す。

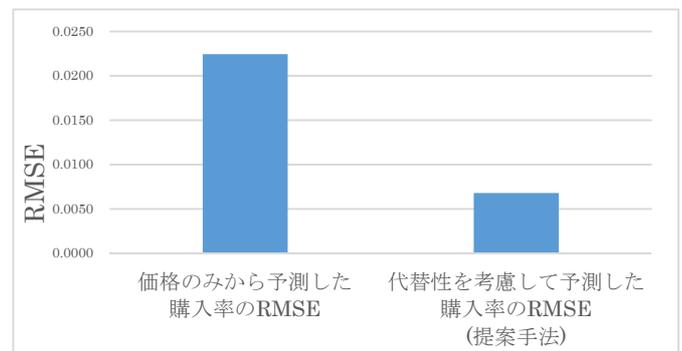


図 3. 購入率予測の実験結果

図 3 から、代替商品の購入率を考慮した場合に予測精度が向上していることがわかる。したがって、今回選択した牛乳 A と牛乳 B の組み合わせについては、代替関係を考慮した推薦手法が適用可能であることを確認した。代替商品の自動抽出については今後の課題として引き続き検討する。

6. まとめ

代替性を考慮して商品の購入率予測を行った場合、商品価格のみで予測した場合に比べて予測精度が向上するという結果が得られた。今後は、他の商品カテゴリについて代替性の検証を進めるとともに、代替商品の自動抽出手法をユーザの商品に対する嗜好度を考慮した推薦手法について検討を行う。

参考文献

- [Sarwar 01] Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl: Item-based Collaborative Filtering Recommendation Algorithms, Proceedings of the 10th international conference on World Wide Web pp.285-295, 2001
- [松村 16] 松村 直樹, 和泉 潔, 山田 健太: POS データに基づく欠品時の顧客行動を考慮した小売店舗の購買シミュレーション, 人工知能学会論文誌 vol.31, 2016
- [高橋 92] 高橋 克也: 食品アイテム間の競合分析 -POS データを用いた実証分析-, 農総研季報第 15 号, 農林水産政策研究所, 1992