F-016

アソシエーションルールを用いた推薦システムにおける評価履歴選択による精度向上に関する検討

A Study on Improvement of Accuracy of in Recommendation System using Association Rule by Selecting Evaluation Histories

伊藤 寛明 † 吉川 大弘 † 古橋 武 †
Hiroaki Ito Tomohiro Yoshikawa Takeshi Furuhashi

1. はじめに

近年，インターネットの普及により電子商取引が増加しており，それに伴い EC サイトでは膨大な数の商品を扱うようになってきている．その結果，それら超多種の商品の中から，ユーザの嗜好にあった商品をユーザ自身で探し出すことが困難となり，推薦システムの利用が期待されている．この推薦システムの代表的な方法に，ユーザの評価履歴をもとに推薦を行う手法である協調フィルタリングがある．一方，EC サイトで扱っているような大量にあるデータの中から，価値のある情報を抽出するデータマイニング手法の一つにアソシエーション分析 [1] がある．これは，関係性の強い組み合わせをアソシエーションルールとして抽出し，新たな知見を得るために用いられる．この手法をユーザの評価履歴に対して適用し，協調フィルタリングによるアイテム推薦に用いた研究が報告されている [2][3].

一方，推薦システムにおいて，ユーザが推薦されたアイテムを好んだ／購入した割合である“精度”は，最も重要な評価指標の 1 つである．しかし近年，ユーザ満足度の観点から，精度に加えて，“意外性”や“説明性”に対する評価の必要性が指摘され始めている [3].

推薦を行う際，推薦回数が増えるにしたがってユーザの評価履歴も増えるが，すべての評価履歴の重要度が等しいわけではない．例えば，「あるアイテム A を好む人の 50%の人がアイテム B を好む」という評価履歴は，アイテム B を推薦する／しないという判断に対して情報量の増加（推薦の精度）に寄与しないと考えられる．そこで本稿では，協調フィルタリングによる推薦システムおいて，適切な指標を用いてユーザの評価履歴を選択して用いることで，推薦システムの精度および意外性の向上を図る．

2. 推薦システム

2.1 アソシエーション分析

アソシエーション分析とは，データの中から価値のある組み合わせ（アソシエーションルール）を見つけ出す手法である．例えば，スーパーマーケットやコンビニエンスストアなどの売り上げデータから，頻繁に購入される商品の組み合わせを見つけ出し，商品の陳列に反映させることなどに応用されている．

アソシエーションルールは，A ⇒ B と表され，A は条件部，B は結論部と呼ばれる．このルールは，A という事象が生じたときに，B という事象が生じるという意味をもつ．代表的なアソシエーションルールの評価指標として $confidence$ がある．

$$confidence = \frac{N(\mathrm{A} \cap \mathrm{B})}{N(\mathrm{A})} \tag{1}$$

$N(\mathrm{A})$ は条件部 A，$N(\mathrm{A} \cap \mathrm{B})$ は条件部 A と結論部 B を同時に満たすデータの件数（本稿においてはユーザ数）である．以下，アソシエーションルールを用いた協調フィルタリングについて説明する．

2.2 アイテムベース協調フィルタリング

推薦を行うユーザ（以降，“対象ユーザ”と呼ぶ）の評価履歴をアソシエーションルールの条件部に用いて，結論部に各アイテムに対する評価「Like」を当てる．例えば，対象ユーザがアイテム A に対して「Like」と評価し，アイテム B が未評価であるとき，全ユーザに対して求められる「アイテム A=Like ⇒ アイテム B=Like」の $confidence$ を，アイテム B のスコアに加算する（アイテム A が「Don't Like」のときは,「アイテム A=Don't Like ⇒ アイテム B=Like」の $confidence$）．対象ユーザのすべての評価履歴により未評価のアイテムのスコアを求め，最もスコアの高いアイテムを推薦する．

2.3 ユーザの評価履歴の選択

本稿では，上述の未評価アイテムのスコアを求める際，すべての評価履歴を用いるのではなく，情報量の大きい評価履歴を選択して用いることで，推薦システムの精度・意外性の向上を図る．ここでは，対象ユーザの評価履歴を選択する際に,「アイテム A を Like と評価した人の 90%もの人がアイテム B を Like と評価している ($confidence$=0.9)」と,「アイテム A を Like と評価した人の 10%しかアイテム B を Like と評価していない ($confidence$=0.1)」から得られる情報は等しく高く，$confidence$=0.5 の場合が得られる情報は最も少ないとみなす.

本稿では，表 1 のようなユーザの評価履歴（1 列目）と推薦候補のアイテム（1 行目）からなる行列を用いて推薦を行う．各要素内の値は，条件部 A と結論部 B から計算される $confidence$ を，括弧内の値は以下により計算される IV の値を表している．

$$IV = \frac{1}{confidence\,(\mathrm{A} \Rightarrow \mathrm{B})} + \frac{1}{confidence\left(\mathrm{A} \Rightarrow \bar{\mathrm{B}}\right)} \tag{2}$$

IV は，その要素の $confidence$ が 0.5 の場合に最小値 (2.0) をとり，$confidence$ が大きくまたは小さくなるほど大きな値となる．

表 1 において，すべての評価履歴を用いた場合は，アイテム B2 が推薦される．しかし，例えば IV の高い上位 2 個を選択した場合，アイテム B2 における「アイテム A2=Don't Like ⇒アイテム B2=Like」の $confidence$

†名古屋大学工学研究科

が 0.05 であるという情報を優先して用いることとなり，アイテム B1 が推薦される．

3. 実験

3.1 使用データ

実験には MovieLens[4] を用いた．映画に対する 5 段階の評点のうち，1 から 3 を「Don't Like」，4 と 5 を「Like」として実験を行った．ただし,「Like」と「Don't Like」をそれぞれ 20 回以上評価したユーザ 436 人，50 人以上に評価された 534 のアイテムを対象とした．

3.2 精度・意外性の評価

2.3 で示した IV の上位 4 個の評価履歴を用いた場合に，最も精度の向上に対して効果があるという結果となった．本実験では，対象ユーザにおいて，ランダムに選択した 10 個の評価済みアイテムに対する評価履歴が与えられた状態から，その他の評価済みアイテムを未評価としてアイテムの推薦を 10 回行った．10-fold cross-validation により実験を行い，10 試行の平均値を求めた．推薦システムの評価指標を以下に示す [3].

$$\frac{1}{N}\sum_{i=1}^{N} t_i \tag{3}$$

a) 精度

推薦回数を N, 推薦アイテムの集合を I=$\{I_1, I_2, ..., I_N\}$, I_i に対する評価履歴を e(I_i) = 1/-1 とすると, 以下の式で表される．

$$t_i = \begin{cases} 1 & if \quad e(I_i) = 1 \\ 0 & otherwise \end{cases} \tag{4}$$

精度は，対象ユーザが推薦されたアイテムに対して「Like」と答えた割合である．

b)Novelty

$$t_i = \begin{cases} 1 & if \quad e(I_i) = 1 \ \ and \ \ I_i \notin I_{NP} \\ 0 & otherwise \end{cases} \tag{5}$$

I_{NP} は Non-Personalized 法における推薦アイテムの集合であり，Novelty は推薦アイテムが「Like」，かつ Non-Personalized な推薦には現れない割合である．

c)Personalizability[3]

$$t_i = \begin{cases} log_2 \frac{1}{P(e(I_i)=1)} & if \quad e(I_i) = 1 \\ 0 & otherwise \end{cases} \tag{6}$$

$P(e(I_i) = 1)$ は，全ユーザにおける推薦アイテムの「Like」割合である．Personalizability は，推薦アイテムの「Like」の割合の低さを情報量にしたもので，推薦されたアイテムが「Like」，かつそのアイテムの「Like」割合が小さいほど大きな値をとる．

図 1 に結果を示す．IV が上位となるもののみを用いることで，すべての評価履歴を用いる場合に比べて，わずかではあるが精度を向上させることができた．また，意外性を示す Novelty や Personalizability に関しても，同様の結果を得ることができた．

表 1: 推薦のために用いる行列

	B1=Like	B2=Like
A1=Like	0.75(5.3)	0.75(5.3)
A2=Don't Like	0.15(7.8)	0.05(21)
A3=Like	0.35(4.4)	0.55(4.0)
A4=Don't Like	0.45(4.0)	0.45(4.0)
全評価履歴	1.7	1.8
IV 上位 2 個	0.9	0.8

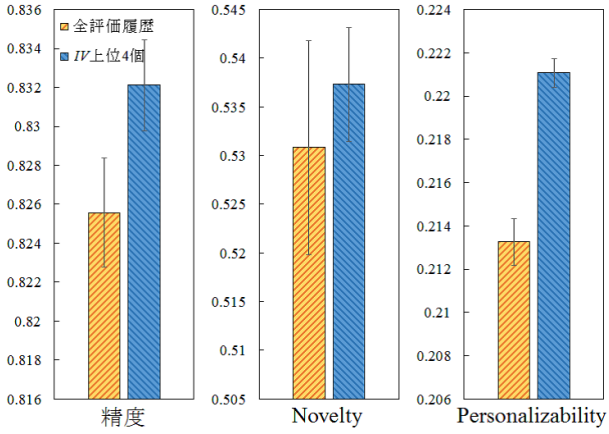

図 1: 全評価履歴と IV 上位 4 個の比較

4. おわりに

本稿では，アソシエーションルールを用いた推薦手法において，情報量の増加に寄与しないと考えられる評価履歴を除くことで，わずかではあるが精度および意外性を向上させることができることを示した．今後は，精度を維持し，意外性を向上させる手法に対する検討を行っていく予定である．

参考文献

[1] Rakesh Agrawal, Ramakrishnan Srikant: Fast Algorithms for Mining Association Rules in Large Databases, 20th International Conference on VLDB, pp.487-499, 1994

[2] Weiyang Lin, Sergio A. Alvarez: Efficient Adaptive-Support Association Rule Mining for Recommender Systems, Data Mining and Knowledge Discovery, Vol.6, No.1, pp.83-105, 2002

[3] 吉川大弘，森貴章，古橋武: Personalizability を考慮した推薦システムの提案，情報処理学会 MPS-90, 2012

[4] Bradley N. Miller, Istvan Albert, Shyong K. Lam, Joseph A. Konstan, and John Riedl: MovieLens unplugged: experiences with an occasionally connected recommender system, 8th international conference on IUI, pp.263-266, 2003