

アソシエーションルールを用いたアイテム推薦における アイテムベースとユーザベースの性能比較

Performance comparison between item-based method and user-based method
using association rule for recommendation of items

伊藤寛明

Hiroaki Ito

吉川大弘

Tomohiro Yoshikawa

古橋武

Takeshi Furuhashi

名古屋大学工学研究科

Graduate School of Engineering Nagoya University

In recent years, with the spread of the electronic commerce via the Internet, recommendation system is practically used in many applications. Collaborative filtering is one of the typical methods in recommendation system. This paper aims to apply association analysis to recommendation which is a method in data mining techniques. In recommendation, "Accuracy" is a significant evaluation index. When we recommend items which a lot of people like, the accuracy is likely to be high. Then it has been said that the accuracy is not the only evaluation index but "Serendipity" is also required in terms of user's satisfaction. We compare the user-based method and the item-based method in collaborative filtering using association analysis in terms of accuracy and serendipity. Based on the consideration of the feature of user-based method and item-based method, we give a basic try to combine them.

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。そのため、そ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協調フィルタリングがある。協調フィルタリングとは、ユーザの評価履歴をもとに推薦を行う手法であり、協調フィルタリングのメモリベースは、ユーザベースとアイテムベースに分けられる [Hijikata 07]。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がある [Agrawal 94]。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。この手法は、ユーザの評価履歴に対しても同様に適用できるため、協調フィルタリングのアイテム推薦に用いることができる [Lin 02][Yoshikawa 12]。

本稿では、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、ユーザベースとアイテムベースの推薦システムに対する比較を行う。また、それにより得られたユーザベースとアイテムベースの特徴から、評価履歴数によってそれらを切り替える推薦システムについて検討する。

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、 $A \Rightarrow B$ とあらわし、 A は条件部、 B は結論部と呼ばれる。このルールは、 A という事象が連絡先: 伊藤寛明, 名古屋大学大学院工学研究科, 名古屋市千種区不老町, 052-789-2793, 052-789-3166, itou@cplx.cse.nagoya-u.ac.jp

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。

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

$N(A)$ は条件部 A , $N(A \cap B)$ は条件部 A と結論部 B を同時に満たすデータの件数である。本稿においては、ユーザ数またはアイテム数となる。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 ユーザベース協調フィルタリング

ユーザベースでは、ユーザの嗜好 (Like/Don't Like) をアソシエーションルールの条件部に用いて、結論部には推薦を行うユーザ (以降、「対象ユーザ」と呼ぶ) の評価「Like」を当てて、例えば、対象ユーザとユーザ 1 の評価履歴に基づいて求められた「ユーザ 1=Like \Rightarrow 対象ユーザ=Like」というルールに対する *confidence* が、対象ユーザとユーザ 1 との類似度をあらわすと考えられる。そのため、もしユーザ 1 が対象ユーザと類似度が高く、アイテム a (対象ユーザの未評価アイテム) に「Like」と評価をしていたら、対象ユーザにとってもアイテム a が「Like」となる可能性が高いとして、*confidence* をアイテム a のスコアに加算する。これは、「ユーザ 1=Don't Like \Rightarrow 対象ユーザ=Like」に対しても計算する。これをすべてのユーザ、すべての対象ユーザの未評価アイテムについて求め、最もスコアの高いアイテムを推薦する [Lin 02]。

2.3 アイテムベース協調フィルタリング

一方、アイテムベースでは、対象ユーザ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」を当てて、例えば、対象ユーザがアイテム 1 に対して「Like」と評価し、アイテム a を未評価であるとき、全ユーザに対して求められる「アイテム 1=Like \Rightarrow アイテム a =Like」の *confidence* を、アイテム a のスコアに加算する (アイテム 1=Don't Like のときは、「アイテム 1=Don't Like \Rightarrow アイテム a =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 [Yoshikawa 12]。

3. 実験

3.1 ユーザベースとアイテムベースの比較実験

3.1.1 使用データ

MovieLens[Miller 03] と JesterJoke[Gupta 99] を用いて実験を行った。MovieLens は、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。JesterJoke は、ジョークに対する -10.00 から 10.00 の評点について、2.00 未満を「Don't Like」、2.00 以上を「Like」として実験を行った。

3.1.2 実験方法

実験の流れを以下に示す。

- 1) 対象ユーザの評価データから、最初のアイテムを推薦する際に用いる評価履歴を指定個数ランダムに選択し、それ以外をすべて未評価アイテムとする。
- 2) 1) で選んだアイテムの評価情報を用いて、2.2, 2.3 で述べた方法により、対象ユーザにアイテムを推薦する。
- 3) 推薦されたアイテムが「Like」か「Don't Like」かを、対象ユーザの元データの評価を用いて判定する。
- 4) 推薦されたアイテムの元データでの評価を評価履歴に加え、指定回数推薦を繰り返す。
- 5) 3) の判定結果に基づき推薦システムを評価する。

MovieLens では、「Like」を 50 回以上、かつ「Don't Like」を 10 回以上評価したユーザ 351 人、20 人以上に評価された 827 のアイテムを対象とした。JesterJoke では、「Like」を 70 回以上かつ「Don't Like」を 20 回以上評価したユーザ 800 人、20 人以上に評価された 100 のアイテムを対象とした。最初のアイテムを推薦するのに用いる評価履歴数として、MovieLens では、{1, 5, 10, 20}、JesterJoke では、{1, 5, 10, 20, 30, 40} と変えて 10-fold cross-validation により評価実験を行った。どちらについても、アイテムの推薦回数を 10 回とした。ユーザベースとアイテムベースを対等に比較するため、各試行においてはこれら 2 つの方法で同じアイテムの評価履歴を用いた。推薦システムの評価指標を以下に示す [Yoshikawa 12]。

$$\frac{1}{N} \sum_{i=1}^N t_i \quad (2)$$

a) 精度

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

$$t_i = \begin{cases} 1 & \text{if } e(I_i) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

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

$$t_i = \begin{cases} 1 & \text{if } e(I_i) = 1 \text{ and } I_i \notin I_{NP} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

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 [Yoshikawa 12]

$$t_i = \begin{cases} \log_2 \frac{1}{P(e(I_i)=1)} & \text{if } e(I_i) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$P(e(I_i) = 1)$ 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。

3.1.3 結果と考察

図 1 に MovieLens、図 2 に JesterJoke の結果を示す。凡例の I はアイテムベース、U はユーザベースをあらわしている。横軸は、最初に用いた評価履歴の数である。

図 1 より、MovieLens に適用した場合、ユーザベースよりもアイテムベースが精度が高いことがわかる。また図 2 の JesterJoke においては、評価履歴数が少ない場合については MovieLens と同様の結果となったが、評価履歴数が多い場合、ユーザベースがアイテムベースを上回る結果となった。一方、意外性をあらわす Novelty と Personalizability は、図 1 の MovieLens、図 2 の JesterJoke 共にユーザベースが高いことがわかる。この結果は一般的にいわれている特徴と等しく、本稿で用いた推薦システムが有効であったと考えられる。また、[Yoshikawa 12] で示されている Personalizability が、意外性をあらわす指標の一つとして適切であることが確認できた。

次に評価履歴数の増加による精度の変化を比べると、図 1 の MovieLens では、アイテムベース、ユーザベース共に変化が見られないことがわかる。一方、図 2 の JesterJoke の精度において、アイテムベースでは変化がないのに対して、ユーザベースでは評価履歴数が多くなるほど上昇し、評価履歴数 20 以上ですべての指標においてアイテムベースを上回る結果となった。データセットの未評価割合である疎性を比べると、MovieLens で 0.774、JesterJoke で 0.001 となり、JesterJoke は密なデータであることが確認できた。ユーザベースでは、データが密な場合、対象ユーザの評価履歴数に比例して対象ユーザと他のユーザ間で共通して評価したアイテムが増える、つまり、十分な評価数からユーザ間の類似性を計算することが可能となるため、精度が上昇したと考えられる。

3.2 ユーザベースとアイテムベースとの切り替え

3.1 より、JesterJoke において、評価履歴数が少ない場合はアイテムベースの精度が高く、評価履歴数が多い場合はユーザベースの精度が高い結果となった。そこで精度に注目し、評価履歴数によってアイテムベースとユーザベースを切り替える実験を行った。アイテムベースからユーザベースへの切り替えの閾値として、対象ユーザと他のユーザ間で共通して評価したアイテム数の平均値を用いた。3.1.1, 3.1.2 と同様のデータおよび手順で実験を行った。

図 3 に、上述の切り替えによる推薦を行った際の (a) 精度、(b) Novelty、(c) Personalizability を示す。凡例のアイテムベース、ユーザベースは、10 回すべてそれぞれの方法で推薦を行った際の結果である。横軸は、アイテムベースからユーザベースに切り替える共通評価アイテム数の閾値となっている。

図 3(a) の精度において、切り替えの閾値を調整することによって、アイテムベース、ユーザベースそれぞれで推薦を行うよりも結果的に高い精度を得られていることがわかる。また図 3(b) の Novelty、図 3(c) の Personalizability については、閾

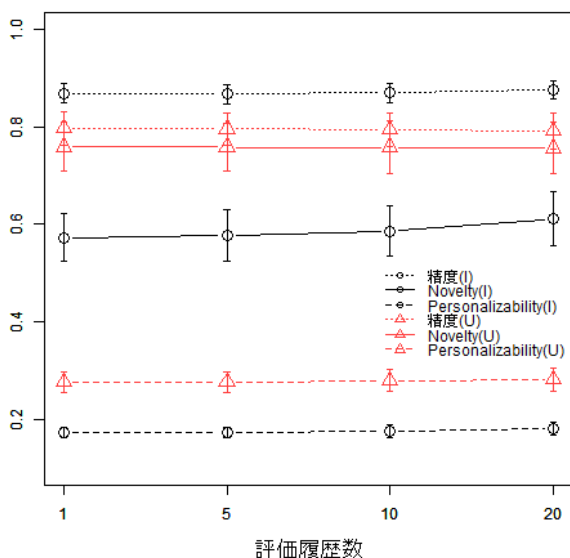


図 1: MovieLens

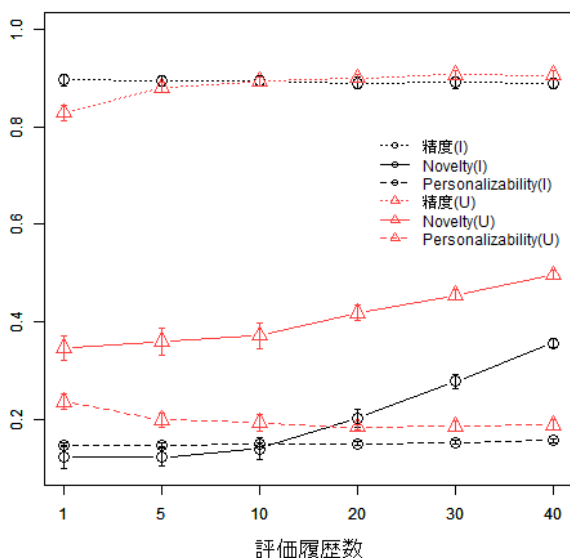


図 2: JesterJoke

値を上げるにつれ、それぞれの値が下がる結果となった。切り替えにより、Novelty, Personalizability において、アイテムベースとユーザベースとの中間値となった。

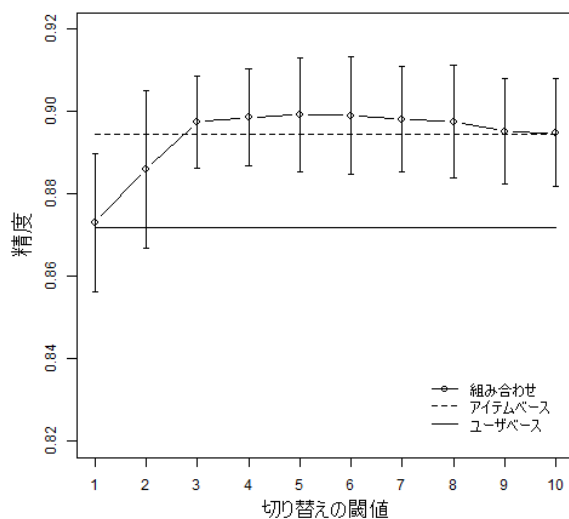
4. おわりに

本稿では、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、ユーザベースとアイテムベースを比較した。アイテムベースにおいては精度が、ユーザベースにおいては、意外性をあらわす Novelty や Personalizability がそれぞれ高

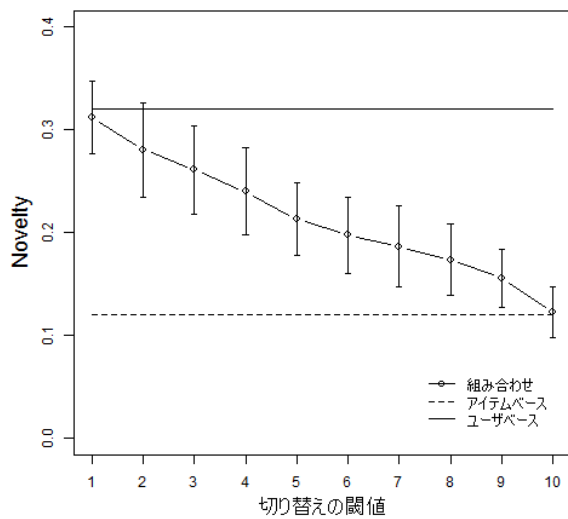
い傾向にあること、さらにデータが密で評価履歴が多い場合には、すべての指標においてユーザベースがアイテムベースを上回ることを確認した。また、これらの特徴から、評価履歴数によってアイテムベースからユーザベースに切り替える推薦方法について検討した。ただし実際の推薦システムでは、疎なデータを扱うことが多いと考えられるため、実用的なデータにおける、アイテムベースに対する意外性の向上、ユーザベースに対する精度の向上、また、それらを組み合わせる方法についての検討が今後の課題である。

参考文献

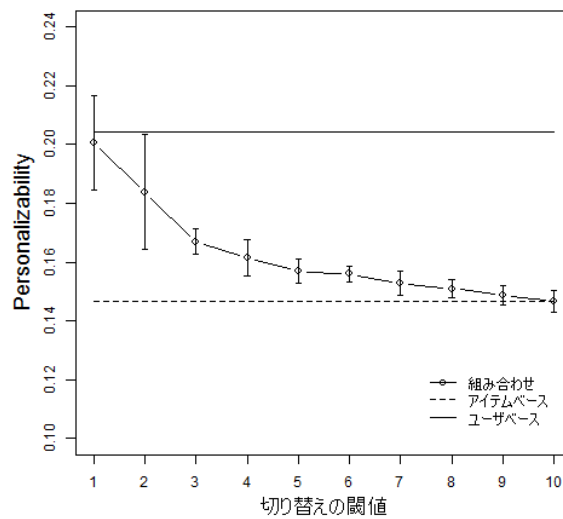
- [Hijikata 07] 土方嘉徳: 嗜好抽出と情報推薦技術, 情報処理学会論文誌, Vol.48, No.9, pp.957-965, 2007
- [Agrawal 94] Rakesh Agrawal, Ramakrishnan Srikant: Fast Algorithms for Mining Association Rules in Large Databases, Proceedings of the 20th International Conference on Very Large Data Bases, pp.487-499, 1994
- [Lin 02] Weiyang 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
- [Yoshikawa 12] 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
- [Miller 03] Bradley N. Miller, Istvan Albert, Shyong K. Lam, Joseph A. Konstan, and John Riedl: MovieLens unplugged: experiences with an occasionally connected recommender system, Proceedings of the 8th international conference on Intelligent user interfaces, pp.263-266, 2003
- [Gupta 99] Dhruv Gupta, Mark Digiovanni, Hiro Narita, and Ken Goldberg: Jester 2.0 (poster abstract): evaluation of a new linear time collaborative filtering algorithm, Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pp.291-292, 1999



(a) 精度



(b) Novelty



(c) Personalizability

図 3: アイテムベースとユーザベースの切り替え