

コスメレビュー共有システムのための類似ユーザ判定手法

奥田 麻美[†] 松波 友稀[†] 上田真由美^{††} 中島 伸介[†]

[†] 京都産業大学コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

^{††} 流通科学大学経済学部 〒651-2188 兵庫県神戸市西区学園西町 3-1

E-mail: [†]{g1444341,g1245108,nakajima}@cc.kyoto-su.ac.jp, ^{††}Mayumi_Ueda@red.umds.ac.jp

あらまし 近年、商品に対するレビュー情報を閲覧可能な口コミサイトが多く存在し、多くの人が商品購入時にレビュー情報を参考にしている。特にコスメアイテムにおいては、直接肌に使用するものであるためユーザが自分の肌に合わないアイテムの購入を避けたいという傾向が他の商品に比べてより顕著であると考えられる。そこで、我々は個々のユーザにとって真に有用なレビュー情報を提供することが可能なコスメレビュー共有システムの構築を目指している。このようなシステムの実現には、価値観の共有を可能とするユーザを発見し、そのユーザによるレビューを推薦することが効果的であると考えられる。そこで、本稿では、価値観の共有が可能となるユーザを発見することを目的に、ユーザ間の類似度を算出する方法を提案する。類似ユーザの判定には、我々の先行研究で実現する評価項目別スコアを用いる。

キーワード コスメ、レビュー、ユーザ間類似度、レビュー推薦、レビュー共有システム

1. はじめに

近年、商品に対するレビュー情報を共有し、閲覧できるユーザ投稿型のクチコミサイトが多く存在しており、多くのユーザがその商品レビューを参考にして商品の購入を行っている。その中でもコスメアイテムにおいては、直接肌に使用するものであるためユーザが自分の肌に合わないアイテムの購入を避けたいという傾向が他の商品に比べてより顕著であると考えられる。コスメアイテムに関するレビューサイトとしては、@cosme [1] や MAQUIA [2] 等が存在し、若い女性を中心に広く利用されている。しかし、これらのサイトを利用しても、本当に自分にあったコスメを見つけることは安易でない。なぜなら、他の投稿ユーザがあるコスメアイテムに対するレビュー情報として、「肌に合う」と記述していたとしても、肌質などの違いにより自分の肌に合うという保証はないからである。コスメアイテムと肌の相性は各個人でそれぞれ異なるため、好みだけでなくコスメアイテムの使用感についても類似するユーザを探し、そのユーザのレビューを参考にしていくという手順が必要になる。すなわち、好みだけでなく、使用感の似たユーザを価値観の共有可能な類似ユーザとして見つけ出し、各ユーザにとって真に有用なレビューを推薦することが可能なレビュー推薦システムを実現する意義は非常に高いと考える。

このようなシステムを実現するためには、各投稿者のコスメアイテムに対するレビューから好みおよび使用感を推定し、ユーザ間の類似度を算出する必要がある。しかし、既存のレビュー提供サービスでは、各アイテムに対して総合的な評価を行うものが多く、細かい価値観を抽出し、類似度を算出することが困難である。そこで、我々は、先行研究でユーザによるレビュー情報を分析し、複数の評価項目に対して得点付けする手法を開発している [8]。本研究では、この評価項目別の得点を用いて、ユーザ間の類似度を算出する手法を提案する。

以下、2章では関連研究との比較を述べる。3章ではコスメレビュー共有システムの構築について述べると共に、4章で類似ユーザ判定手法について述べる。最後に5章でまとめと今後の課題を述べる。

2. 関連研究

レビュー情報を扱うサイトがインターネット上に多く存在している。Amazon.com [3] や価格.com [4] などさまざまな商品を取り扱うサイトや、レストラン情報を提供する食べログ [5] などが有名である。また、近年では美容系のレビュー情報を扱うサイトも多く存在しており、美容院やサロンなどの情報を取り扱うホットペーパービューティー [6] やコスメ・美容情報を取り扱う@cosme などが有名である。@cosme の運営会社によると、2016年10月の時点で月間2.8億ページビュー、メンバー数が3,900,000人、総クチコミ件数が約13,000,000件と報告されており [7]、多くのユーザに利用されていることがわかる。

レビュー情報を扱うサイトが広く普及したことで、レビューを分析する研究が多く行なわれている。我々の先行研究でも、@cosme から収集した化粧水に関するレビューを分析し、コスメアイテムに対する評価表現辞書の作成を行い、それを用いた評価項目別レビューの自動スコアリング方式の開発を行っている [8]。

中辻らは、アイテムの内容をアイテムに与えられているメタデータをもとに表現している。さらに、ユーザの記述したレビューを用い、アイテムのメタデータの属性値に対して sentimental analysis を行っている。アイテムだけでなく与えられたメタデータに対し形容表現が関連付けられたか否かの分析を行うことで、ユーザがアイテムに対しどのような感性を持って評価を下しているかを解析している [9]。アイテムに対する評価のみでなく、アイテムに対し与えた形容表現も共通するユーザを類似ユーザとして取り扱えるため、多様なクラスに所属する

アイテムを精度よくユーザに推薦できると述べている。本研究では、コスメアイテムを対象としており、アイテムそのものの評価により好みに合うか否かを推定するのみでなく、アイテムの使用感がユーザ自身にとって適切か否かについても推定したうえで、類似ユーザの判定を行う。

古澤らは、サービス選定に QoS (Quality of Service) に加えて過去に使用したサービスの評価値を用いることで、主観的、客観的両方の品質を考慮した Web サービスの推薦を行うことを目的とし、ユーザの嗜好に合わせた推薦システムを提案、検証している [10]。この取り組みでは、類似ユーザの抽出に UDDI リポジトリに蓄えられたサービスデータ (QoS) と評判リポジトリに蓄積された評判データから、協調フィルタリングにより処理を行い、推薦対象とする Web サービス群を作成している。本研究では、コスメアイテムに対する評価項目別の評価値を用いることで、ユーザの嗜好にあったアイテムの提案のみでなく、アイテムの使用感の似た類似ユーザの判定を行う。

中辻らは、被推薦ユーザがまだ評価した事の無いドメインのアイテムを精度良く推薦可能とする手法を提案している [11]。ユーザのアイテムへの評価値をそのアイテムを含むクラスへの評価値として反映させることにより、アイテムのみでなくクラスへの評価の一致性も用いてユーザ間の類似度を計算できる。それにより、ユーザの所持するアイテム数が少ない場合においてもユーザ間の類似度を適切に計算できると述べている。また、ドメイン特化ユーザグラフ (Domain-specific-user graph: DSUG) を構築する際のユーザ類似度計算法として、アイテムの一致性のみに基づく手法よりも、タクソノミに基づく手法の方が、精度が大きく向上することも確認している。本研究では、評価項目別スコアの平均値によるコスメアイテムのクラスタ分類を行い、同一アイテム群のあるアイテムに対して似た評価をするユーザ同士を類似ユーザと判定することでデータが少ないという問題を解決する。

伊藤らは、全ユーザの嗜好傾向から書くユーザの相対的な嗜好を把握する手法を用い、推薦制度の向上を試みている [12]。主成分分析を行うデータについて、推薦対象ユーザのアイテムに対する評価データだけでなく、全ユーザの全アイテム評価を用いることで、全員の評価傾向の中での対象ユーザの特徴を推測することにより精度の向上を図っている。

伊藤らは、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・検討を行っている [13]。基本的には、アイテムベースが精度が高いこと、データが密な場合には、すべての指標においてユーザベースが優れていることを確認している。

3. コスメレビュー共有システムの構築

本研究の最終的な目標は、好みおよび使用感において真に有益なレビュー情報を効率的に推薦することが可能なレビュー推薦システムの構築である。本稿では、その一部として、好みおよび使用感において類似するユーザを判定する手法について提案する。本章では、3.1 節でコスメレビュー共有システムの概要を述べ、3.2 節で評価項目別自動スコアリング手法について

述べる。

3.1 コスメレビュー共有システムの概要

図 1 は、将来的に構築を目指すコスメレビュー共有システムの概要図である。このような推薦システムを実現することにより、対象ユーザは評価項目別のスコアを見ることができ、好み度および使用感が類似するユーザに基づいた真に有益なレビュー情報のみを効率的に得ることができる。本稿では、図 1 の赤字 ((a)~(e)) で示した部分について述べる。レビュー推薦システムは、以下の手順で動作する。

- あるアイテムに関するレビューについてユーザが問合せ。
- (a) にて問合せを行った対象ユーザとの類似ユーザを類似ユーザ抽出モジュールに問合せ。
- 評価値付きレビュー DB から、レビューデータおよびレビュー投稿ユーザ情報を取得し、対象ユーザとの類似ユーザを特定。
- (c) にて特定されたに類似ユーザによるレビューを、レビュー推薦モジュールへ抽出。
- 対象ユーザへ類似ユーザに絞ったレビューを推薦。

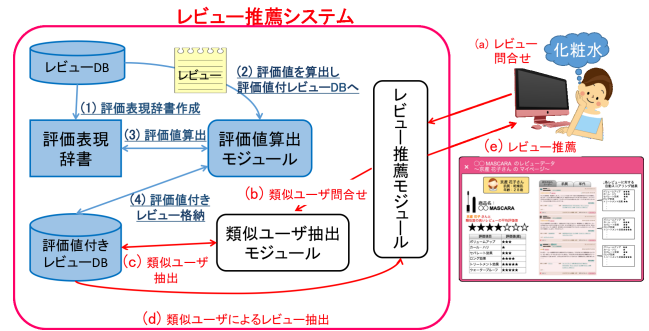


図 1 レビュー推薦システムの概要

3.2 評価項目別自動スコアリング

本研究の目的である類似ユーザの判定手法では、アイテムに対する評価項目別スコアが必要となるが、従来のレビューサイトでは、各アイテムに対する総合評価としてスコアが付与されていることが多い。そこで、本研究では先行研究 [8] で取り組んでいる評価項目別の得点を用いる。本節では、評価項目別自動スコアリングの概要を述べる。

図 1 における、青字 ((1)~(4)) で記載した処理がレビューの評価項目別自動スコアリングの処理内容である。以下に処理の手順を示す。

- レビュー DB から抽出したレビューデータを分析し、評価表現辞書を作成。
 - レビュー DB から評価値を算出するためのレビューを抽出。
 - (1) にて作成した評価表現辞書を用いて、(2) で収集したレビューの評価値を算出 (自動スコアリング)。
 - (3) にて評価値が付与されたレビューを評価値付きレビュー DB に格納。
- (3) の自動スコアリングでは、抽出したレビューに対してコスメアイテムの評価軸を表すキーワードの有無を調べ、存在した

場合に同一文中に共起する特徴語及び程度を表す語の有無を確認する。その後、キーワードの共起に基づく評価表現辞書に問合せ、該当するスコアを取得する。上記のことを各評価項目別に集計し、自動スコアリングを実現する。

4. 類似ユーザ判定手法

本章では、類似ユーザ判定手法に関して述べる。まず、4.1節で提案手法の概要を述べ、4.2節でコスメアイテムの分類、4.3節で類似ユーザ判定について述べる。

4.1 提案手法の概要

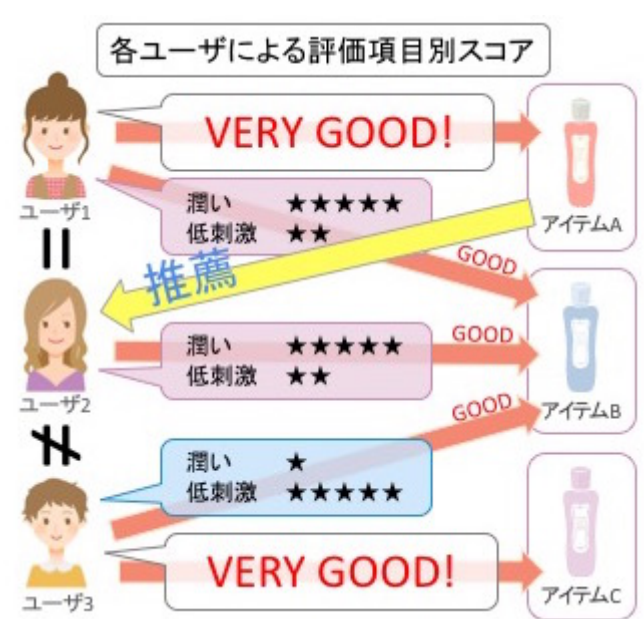


図2 類似ユーザ判定手法を用いたコスメアイテム推薦の概要

我々は、真に有用なレビューを推薦するためには、価値観を共有しうるユーザによるレビューを発見することが重要と考える。そこで本稿では、先行研究[8]により得点付けをした評価項目別スコアを用いて、類似ユーザを判定する手法の提案を目的とする。

図2に概要図を示す。この図では、ユーザ1、ユーザ2、ユーザ3がアイテムBに対して「Good」と評価している。また、ユーザ1がアイテムAに対して、ユーザ3がアイテムCに対して「Very Good!」と評価している。ここで、ユーザ2に対して新規のアイテムを推薦する場合、既存の手法では、ユーザ1が「Very Good」と判断しているアイテムAと、ユーザ3が「Very Good」と判断しているアイテムCを同等に扱うことが考えられるが、我々の手法では、評価項目別のスコアを考慮し、ユーザ1とユーザ2を類似ユーザと判定し、ユーザ1が「Very Good」と判断しているアイテムAを推薦することとする。ユーザ3は総合評価の観点からはユーザ2と類似しているが、評価項目別に見た際に評価が異なるため、価値観を共有することができないユーザであると判断し、アイテムCは推薦対象としない。

4.2 コスメアイテムの分類

本節では、コスメアイテムの分類について述べる。一般的に、類似ユーザとは、同一アイテムに対して似た評価をするユーザ同士のことであると考えられる。しかし、コスメアイテムには膨大な数が存在するため、共通のコスメを使用し、似た評価を行っているユーザを見つけ出すことは非常に困難である。そこで、評価項目別スコアの平均値によるコスメアイテムの分類を行い、同一アイテム群に対して似た評価をするユーザを類似ユーザと判定することでデータが少ないという問題を解決する。

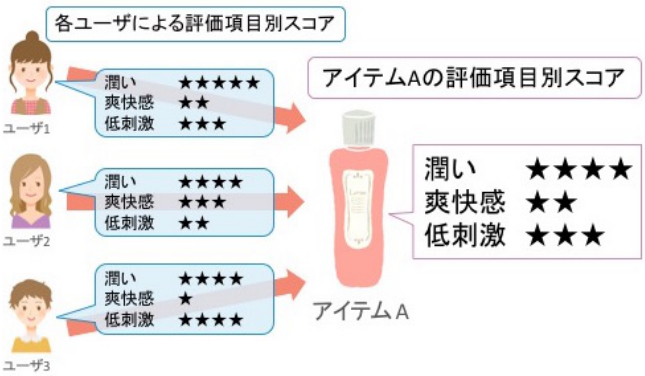


図3 コスメアイテムの評価項目別スコア例



図4 コスメアイテムのクラスタ分類例

以下の手順によってコスメアイテムの分類を行う。

- (1) あるアイテムに対する全ユーザからの評価項目別スコアの平均値をそれぞれ求める。
 - (2) 平均値の似たアイテムを集め、クラスタを作成する。
- 各手順の詳細を以下に述べる。(1) あるアイテムのレビューに対して、評価項目別自動スコアリングによりスコアを付与する。それら全ての評価項目別スコアの平均値を評価項目ごとに求める。つまり、アイテムに対してレビューを行った全ユーザの評価項目別スコアそれぞれの平均値を得ることができる(図3)。(2)(1)で求めた平均値を用い、似た評価のアイテムクラスタ群を作成する。例えば、図4では、クラスタ1は潤い/浸透、美白/UV、低刺激の項目の評価が高く、毛穴・角質ケア、爽快感の項目があまり高くないアイテムのクラスタになっている。また、クラスタ2は、毛穴・角質ケア、爽快感の項目の評価が高く、潤い/浸透、低刺激の項目の評価が低いアイテムのクラスタとなっている。なお、コスメアイテムのクラスタ分類には、

クラスタリング手法の1つである k-means 法を用いることとする。k-means 法は、全データ間の距離を計算する必要がなく計算量も軽く済むため、コスメアイテムの評価項目別平均スコアといった膨大な量のデータを分析するのに有効である。

4.3 類似ユーザ判定

本節では、類似ユーザの判定手法について述べる。前節で述べたコスメアイテムの分類手法によって分けられた、あるアイテム群に対する評価の平均値をユーザ、評価項目ごとに参照する。類似ユーザの判定はピアソンの相関係数を用いて行う。

通常、ユーザが各アイテムに対して評価値を付与しているケースで協調フィルタリングを行う際には、ユーザ数 n 、アイテム数 m となり、 $n \times m$ の評価行列を用いて、類似度を算出している。しかし、提案手法では、各アイテムに対して複数の評価項目に対する評価値がそれぞれ付与されている。すなわち、ユーザ数 n 、アイテム数 m 、1つのアイテムに対する評価項目数 l である時、 $n \times p$ の評価行列を用いる。ただし、 $p = m \times l$ とする。例えば、一般的な手法において、3人のユーザが5アイテムに対して評価を行った場合、3行5列の評価行列となるのに対し、我々の提案手法では、各アイテムに対して10つの評価項目に関して評価値を付与しているため、3行(5×10)列の評価行列となる。ただし、評価値は0~7の数値尺度をとる。本手法により、アイテムに対する単一の評価値の相関ではなく、各アイテムの複数評価項目に対する評価値の相関を確認することができるため、従来の方法よりも精度の高い類似ユーザの判定が可能であると考ええる。

それぞれのアイテム群に対する各評価項目10個

ユーザ	アイテム群1							...	アイテム群5				
	1-1	1-2	1-3	...	1-9	1-10	...	5-1	...	5-8	5-9	5-10	
A	5	4	-	...	7	3	...	-	...	2	2	5	
B	4	-	-	...	5	5	...	7	...	1	2	3	
C	2	2	6	...	-	6	...	4	...	5	-	3	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	

図5 各ユーザのアイテム1~5に対する評価項目別スコア

また、ピアソン尺度ではユーザらの評価値の平均を考慮して計算を行うことによりユーザの比較が可能である。すなわち、ユーザ同士で評価値の絶対値が異なっていたとしても、その評価値に線形な相関があるかないかの確認をすることによって類似ユーザの特定が可能である。そのため、各ユーザによって表現に差が生じるレビューからの評価値を用いても、多くのユーザから類似するユーザを特定することが可能であると考えられる。

図5は、各ユーザのアイテム5つそれぞれに対する評価項目別スコアの例である。アイテム数、評価項目については未定であるが、今回は「化粧品」を対象としており、先行研究[8]に従って評価項目は10項目とした。また、4.2節のクラスタ分類の手法に基づき、化粧品を5個のアイテム群に分類している。

5. おわりに

本稿では、コスメアイテムのレビューサイトにおいて、個々

のユーザに対して真に有用なレビューを推薦するためには価値観を共有可能なユーザを発見することが重要と考え、ユーザの類似度を算出する手法を提案した。具体的には、まず、既存のレビューサイトから得たコスメアイテムに関するレビューを分析して、評価項目別にレビューの自動スコアリングを行ったものを使用した。次に、この評価項目別のスコアを用いて、類似ユーザの判定を行う手法を提案した。提案手法は一般的な協調フィルタリング手法と異なり、各アイテムに対して単一の評価値を持つのではなく、各アイテムに対して複数の評価値を持つ特徴がある。提案手法では、各アイテムに対して複数の評価項目に対する評価値を持ち、その相関をとるため、従来の方法よりも精度の高い類似ユーザの判定、さらには精度の高いレビューを推薦することが可能となると考える。

今後は、複数のデータを用いて評価実験を行い、提案手法で用いたクラスタリング手法の有効性および類似ユーザ判定手法の有効性を検討をしていく。また、問題点としてユーザ同士の評価表現が同様であっても肌質の違いによって意味が変わってくる可能性が考えられる。そのため、評価項目別スコア間での相関を見るなどの手法を検討していく。

謝 辞

本研究で用いたレビューデータは株式会社アイスタイル様より提供いただいた。また、評価項目の検討において、多大なるご助言をいただいたことを感謝します。

本研究の一部は、JSPS 科研費 16K00425 および 26330351 による。ここに記して謝意を表す。

文 献

[1] @cosme, <http://www.cosme.net>

[2] MAQUIA, <https://maquia.hpplus.jp/cosme/>

[3] Amazon.com, <https://www.amazon.com>

[4] 価格.com, <http://kakaku.com>

[5] 食べログ, <https://tabelog.com/>

[6] ホットペッパービューティー, <https://beauty.hotpepper.jp/>

[7] @cosme サイトデータ資料<2016年10月>, 株式会社アイスタイル, <http://www.istyle.co.jp/business/uploads/sitedata.pdf>, (2017年1月4日アクセス)

[8] 松波 友稀, 上田 真由美, 中島 伸介, 階上 猛, 岩崎 素直, John O'Donovan, Byungkyu Kang, コスメアイテム評価表現辞書を用いた評価項目別レビュー自動スコアリング方式, DEIM Forum 2016 論文集, B1-1, 2016.

[9] 中辻 真, 近藤 光正, 田中 明通, 内山 匡, アイテムに係る形容表現を用いたユーザ類似度測定, 人工知能学会全国大会論文集, 3C4-02, 2010.

[10] 古澤 悠, 杉木 優太, 菱山 玲子, 複数ユーザの評判を考慮したWeb サービス推薦エージェントの提案人工知能学会全国大会論文集, 2E2-1, 2011.

[11] 中辻 真, 藤原 靖宏, 内山 俊郎, ユーザグラフ上のランダムウォークに基づくクロスドメイン推薦, 人工知能学会全国大会論文集, 27 巻 5 号 D, 2012.

[12] 伊藤 ゆかり, 波多野 賢治, 松本 尚宏, 明示的な嗜好評価を活用した商品推薦システムの実装とその評価, 情報処理学会第 73 回全国大会, 4M-1, 2011.

[13] 伊藤 寛明, 吉川 大弘, 古橋 武, アイテム推薦のためのアソシエーションルールを用いた類似ユーザの抽出に関する基礎的検討, 情報処理学会第 75 回全国大会, 2P-8, 2013.