

レビューの分析に基づく化粧品推薦方法の提案

日大生産工 ○佐藤 美里 関 亜紀子

1. まえがき

昨今、インターネット上の口コミが話題となり、口コミは消費者の購買行動を左右する重要な要素となり始めている。しかし、個々人の価値観や考え方などは多様であり、特に化粧品の場合は、肌に合う・合わないなどの個人差が問題となる。口コミだけでは口コミを発信した人の肌質が自分と似ているのか、そうではないのかという点について判断できない。

こうした問題に対し、文献1)では、類似ユーザに絞った際の平均評価値をレートで表示する方法を提案している。ここでは各コスメアイテムに対する評価項目毎のレーティングと自由記述により口コミを記述させることで、該当コスメアイテムの平均評価値を計算している。しかしこうした事前に定めた評価項目を用いた方法では、評価項目以外の各アイテムの細かな特徴やユーザの価値観の違いを十分に反映させることができないという課題が残っている。

そこで、本稿ではコスメアイテムに対しての口コミからアイテムの特徴を抽出する。また、個々のユーザの口コミからアイテムに対する不満に相当する項目を抽出する。そして、これらを基に各アイテムの特徴を可視化することで不満を解消するような新しい商品を提案することを目的としている。

2. 評価語や価値観の抽出手順

図1にアイテムの評価語とユーザの価値観の抽出手順を示す。コスメアイテムの評価語の分析では、1つのコスメアイテムについてレビューを解析し特徴(不満)を見つける。同じカテゴリのコスメアイテムでもアイテムごとに違った特徴があり、また1つのコスメアイテムについても良い評価をする者と悪い評価をする者がいる。そこで、カテゴリごと

に評価項目を10個求める。次に、ユーザの1つのカテゴリに対しての口コミを集め、そのユーザが何に重きを置いているかを見つけ出す。1つのカテゴリのアイテムに対して“価格”や“保湿力”の単語が多ければ、そのユーザは“価格”“保湿力”に重きを置いていることになる。

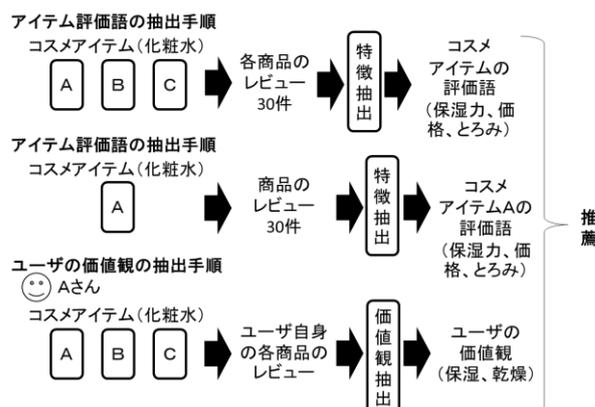


図1 評価語と価値観の抽出手順

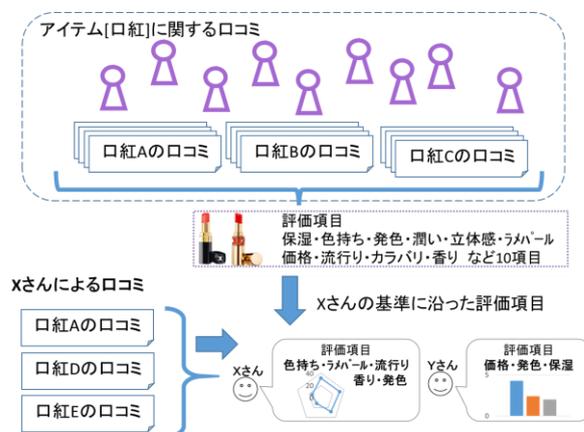


図2 アイテムの可視化と推薦

アイテムの可視化と推薦では、コスメアイテムの評価語の分析で得た評価項目10個の中から、ユーザのタイプに合った評価項目を選択し各商品の特徴の可視化と推薦を行う。例えば口紅の評価項目として“保湿”“色持ち”“発色”“潤い”“立体感”“ラムパール”“価格”

Suggestion of the Cosmetics Recommendation Method Based on User Review

Misato SATOH, Akiko SEKI

“流行り” “カラーバリエーション” “香り”があり、ユーザタイプで得た価値観が価格、発色、保湿の場合は、この3項目に関する各商品の位置づけを可視化し、該当する項目に対する評価値の高い商品を推薦する。

3. コスメアイテムの特徴

アットコスメより化粧水と口紅について5種類ずつ選び、1種類につき30件の口コミを集める。口コミデータは、Khcoderで形態素に分割し、品詞ごとの抽出語リストを作成する。次に、得られた単語に付随する意味を調べるため、単語の共起関係をRMeCabを利用して求める。ここでは、「名詞」「副詞」「形容詞」について、出現頻度が3以上の単語を抽出し出現頻度の高い単語の共起関係を調べる。

Term	Before	After	Span	Total	T	MI
ある	0	9	9	23	2.799763	3.905179
が	0	3	3	125	-0.15285	-0.12201
けれど	0	1	1	1	0.973882	5.258816
た	2	0	2	84	-0.13711	-0.1335
だけ	1	0	1	1	0.973882	5.258816
ちょっと	1	0	1	2	0.947764	4.258816
です	2	0	2	65	0.213783	0.236448
と	0	1	1	50	-0.3059	-0.38504
とても	1	0	1	7	0.817175	2.451461
とろみ	11	0	11	11	NA	NA
の	1	7	8	82	2.071232	1.901264
のに	1	0	1	2	0.947764	4.258816
は	1	0	1	49	-0.27978	-0.35589
ます	2	0	2	92	-0.28486	-0.26475
やや	1	0	1	1	0.973882	5.258816
を	1	1	2	31	0.841701	1.30462
テクスチャー	1	6	7	9	2.556906	4.896246

表1 「とろみ」共起関係

$$T \text{ 値} = \frac{\text{実測値} - \text{期待値}}{\text{実測値の平方根}} \geq 1.6 \quad \dots \text{式(1)}$$

$$MI \text{ 値} = \log \frac{\text{共起回数}}{\text{共起語の期待値}} \quad \dots \text{式(2)}$$

表1に、評価項目「とろみ」に関する共起関係の一例を示す。ここでT値は式(1)により得られる共起関係の有無を調べる指標であり、絶対値が1.65以上であれば、この2つのタームの共起は偶然でないことを示す。MI値は式(2)により求められる指標であり、2つのタームの独立性を示す。この指標が大きい場合、2つのタームは独立ではなく一方のタームが出現すればその共起語が現れる可能性が高いことになる。T値が1.65以上でMI値が高い語を見つけることにより、コスメアイテムの特徴を絞り込める。表1より、「とろみ」「の」と「ある」「テクスチャー」が共起関係にあることが分かる。

4. 評価語と価値観の抽出例

コスメアイテム“化粧水”に対し、コスメアイテムの評価語の抽出とユーザの価値観の抽出を行った結果を表2表3に示す。

表2 コスメアイテムの評価語の抽出結果

アイテム	特徴
化粧水1	潤う・浸透・濃厚・乾燥・とろみ・ハリ感・美白
化粧水2	ニキビに良い・薬用・さっぱり・刺激
化粧水3	浸透・ビタミン・とろみ・香り・ハリ
化粧水4	乾燥・痒み抑える・敏感・潤う・しっとり
化粧水5	しっとり・安い・乾燥・容量

表3 ユーザの価値観の抽出結果

ユーザ	特徴
ユーザ1	効果・乾燥・湿る・香り
ユーザ2	乾燥・効果・とろみ・毛穴

これらの結果は、各化粧水と各ユーザの口コミデータを出現頻度が3以上の単語を抽出し出現頻度の高い単語の共起関係を調べたものであり、共起関係を調べた単語の名詞と動詞だけを抜き出したものである。

1つのコスメアイテムの30件のレビューの中で頻度の高い語はそんなに多くない事が分かった。「すき」「きれい」や「感じ」「使用」などは多く出現するが、アイテムの評価項目にできる「ハリ」「しっとり」「さっぱり」などは出現頻度が低く、また個人的な言い回しが多く検索に引っかからないが多かった。特に「香り」に関しては「やさしい」「さわやか」「独特」「フルーツ」などと一緒に表記されることが多く、「香り」の評価は検討したい。

5. まとめ

出現頻度の高い語の共起関係を調べることでコスメアイテムの特徴を絞り込むことができた。また、アイテム間の優位関係を口コミの数で決めて良いのかも検討課題となる。

参考文献

- 1) 松波友稀 上田真由美 中島伸介/コスメアイテムの使用感および嗜好度判定を目的としたレビュー分析手法の提案