

商品比較のための商品レビュー可視化の一手法

板垣 紫音[†] 上野 史^{††} 太田 学^{††}

[†] 岡山大学工学部情報系学科 〒700-8530 岡山市北区津島中 3-1-1

^{††} 岡山大学学術研究院自然科学学域 〒700-8530 岡山市北区津島中 3-1-1

E-mail: [†]p3w53x7t@s.okayama-u.ac.jp, ^{††}{uwano, ohta}@okayama-u.ac.jp

あらまし EC サイトでは商品の比較を容易にするために、いくつかの評価項目、例えば掃除機の商品における「吸引力」などに対する評判が数値化されている。これらの評価項目は、通常 EC サイトの商品ごとに決まっており変更することができないため、比較したい評価項目が不足している場合がある。そこで本稿では、評価項目を商品レビューから自動生成して商品を比較するための可視化手法を提案する。提案手法では、基準商品と比較対象商品の商品レビューを利用して、商品を比較するための評価項目を自動生成する。そして、自動生成された評価項目を軸とし、その評判を数値化したものを軸の値とするレーダーチャートによって、評価項目の評判を可視化する。ユーザは生成した評価項目に関連する評価項目の評判も比較することができる。大学生と大学院生あわせて 10 名を被験者にして実験を実施し、生成した評価項目の数や妥当性を評価した。その結果、商品にとって重要な評価項目を生成できたが、生成できなかった評価項目もあった。また、被験者に既存の商品比較サイトである価格.com と提案手法によって 2 商品を比較させたところ、全体的には価格.com に及ばなかったが、提案手法で可視化された評価項目を優れていると評価した被験者もいた。

キーワード 可視化, 商品レビュー分析, 評判情報

1 はじめに

近年 EC サイトが広く普及し、インターネット上で商品を購入する機会が増えている。EC サイト上にはユーザによる商品レビューを含む評判情報が多く投稿されている。この評判情報のおかげで、商品を実際に目で見たり、手で触れたりしなくても商品についての情報を得ることができる。評判情報にはユーザが使用した感想やその商品の評価を文章にした商品レビュー、商品の満足度を 5 段階評価で示した情報などが含まれる。また、商品の評価項目、例えば掃除機の「吸引力」や「使いやすさ」などの満足度の評価も存在する。EC サイトでは通常この評価項目は商品ごとに決まっており、変更することができない。例えば、楽天市場¹における商品ジャンル「掃除機」の評価項目は 2023 年 2 月現在では「デザイン」「手入れのしやすさ」「サイズ」「静音性」「パワー」「使いやすさ」の 6 つとなっており変更することはできない。そのため、もしユーザが掃除機の「充電」についての評判を比較したくてもその評価項目は存在しない。つまりこれは、既定の評価項目では不十分である場合といえる。また、同じ「掃除機」のジャンルでも Amazon.co.jp においては商品の評価項目が一部異なっており、また、商品ごとの独自の評価項目もある。しかし、商品ジャンルにおいて共通している評価項目が少なく、購入に際して比較したい評価項目が不足している場合もある。また、これらの評価項目の評判は関連する評判情報を含んでいる。そのため、より具体的な評価項目の評判を得ることができないこともある。例えば、楽天市場の場合、掃除機の「デザイン」について「デザインがインテリア

に合うか」や「持ち手などの細かい部分のデザインの良さ」などの具体的な評価が知りたくても、既定の評価項目である「デザイン」の中からユーザが探すしかない。

そこで本稿では、商品レビューから評価項目を自動で生成し、ユーザが比較したい評価項目だけでなく、それに関連する評価項目で商品を比較できる可視化手法を提案する。提案手法では、ユーザは商品間で共通する評価項目を軸としたレーダーチャートによって商品を比較する。評価項目は主評価項目とその関連する評価項目である副評価項目で構成され、ユーザはレーダーチャートの副評価項目を任意に選択して、主評価項目と表示を切り替えることができる。そのため、評判を比較したい評価項目によって商品を比較できる。また、それぞれの商品独自の評価項目とその評判を数値化したものも同時に可視化する。これにより、ユーザは商品独自の評価項目とその評判についても確認できる。

本稿の構成は以下の通りである。2 節では関連研究について述べる。3 節では提案する商品レビューの可視化手法について述べ、4 節では提案する可視化手法の評価実験について述べる。最後に 5 節で本稿をまとめる。

2 関連研究

2.1 商品の評価項目の抽出に関する研究

藤本ら [1] は評判情報を用いて代替品を推薦する手法を提案した。藤本らの手法では、まず Google Natural Language API によって、商品レビューのテキストから感情表現の対象となる単語とその感情極性を数値化した感情スコアを抽出した。そして、Word2Vec によって抽出した単語の分散表現を得た。その

¹ : <https://www.rakuten.co.jp/>

後、*k*-means 法によって単語をクラスタリングし、得られた単語クラスタをその商品の評価項目とした。

本研究では藤本らと同様に商品レビューから Word2Vec によって単語の分散表現を得る。しかし、本研究では、商品の評価項目として主評価項目と副評価項目を抽出するために、得られた単語のクラスタリングには Ward 法による階層型クラスタリングを用いる。主評価項目に関連する評価項目である副評価項目を抽出することで、商品をより多くの評価項目で比較できる。

平山ら [2] は、レビュー集合から抽出した評価項目に関する好評や不評、有益な文章、他の評価項目との関係を抽出し可視化するシステムを提案した。平山らは商品のレビュー集合を Cabocha で係り受け解析した。そして、係り先が形容詞などの特定の品詞である場合の係り元の名詞を取り出した。その取り出した名詞の内、出現頻度の高いものを評価項目とした。そのため、イヤホンにおける「音」や「音質」というほとんど同義な単語も異なる評価項目となる。本研究では、単語の分散表現をクラスタリングすることにより、類似した単語を一つのクラスタとし、主評価項目を生成する。また、同じクラスタに分類される重要な単語なども抽出し副評価項目とする。そのため同義な単語をまとめて主評価項目として評判を確認することができ、同様に関連する評価項目として副評価項目の評判も確認できる。

2.2 商品比較に関する研究

矢部ら [3] は購入済みのコスメと購入を検討しているコスメの相違点を可視化するシステムを提案した。提案システムではまず、評価表現辞書を用いて各コスメの商品レビューを評価項目別にスコアリングする。次に、ユーザが持っているコスメを元に、理想の評価項目別のスコアを設定させる。提案システムではその理想のスコアに近いコスメをオススメのコスメとしてランキング形式で表示する。オススメのコスメの商品情報等も一緒に表示して、ユーザが比較しやすいようにしている。またユーザは表示されたコスメを選択することにより、理想のスコアを基準として選択したコスメの評価を棒グラフによって確認できる。

一方本稿では、生成した主評価項目と副評価項目の評判をレーダーチャート等で可視化する。それにより、商品をより詳細に比較できる。また、矢部らの手法ではユーザは理想のスコアと実際の商品のスコアを比較するが、本稿ではユーザは実際の 2 商品を比較する。

3 商品比較のための商品レビューの可視化手法

3.1 可視化手法の概要

提案する商品レビュー可視化手法のユーザ-インタフェースの実装には、Python のライブラリである Tkinter²を用いる。

提案手法ではまず、比較する 2 商品のレビューを Google

Natural Language API³に渡し、感情極性分析の対象となる単語とその感情極性を数値化したもの、また文章における重要度を得る。次に獲得した単語を階層的クラスタリングし、階層的なクラスタを生成する。生成したクラスタにおいて、単語の重要度とレビュー内での出現回数を用いてクラスタを代表する単語を選出する。そして、選出した単語を商品比較のための主評価項目とする。また、クラスタ内にある重要な単語をいくつか抽出する。その単語を主評価項目に関連する評価項目の副評価項目とする。次に生成した各クラスタ内の単語の感情極性の値を用いて、主評価項目と副評価項目の評判を評判値としてそれぞれ数値化する。そして、比較したい 2 商品に共通している主評価項目を軸、その評判値を軸の値とするレーダーチャートで比較したい 2 商品の評判を可視化する。また、共通ではない評価項目はそれぞれの商品独自の評価項目としてその評判値とともに可視化する。

3.2 インターフェース

提案する商品レビュー可視化手法のユーザインタフェースのイメージを図 1 に示す。

共通評判情報表示部では、比較する 2 商品に共通する評価項目を軸とし、その評判値を値としたレーダーチャートを表示する。レーダーチャートは青のグラフが左の商品、赤のグラフが右の商品についての評判を示している。また、各軸が主評価項目を表している。マウスポインタを主評価項目のボタンの上に移動すると副評価項目が表示される。主評価項目のボタンをクリックすることにより、主評価項目に関連する副評価項目の評判が表示される。副評価項目の評判の表示の仕方は 2 通りある。2 商品に共通する副評価項目が 3 つ以上ある場合は副評価項目を軸、その評判値を軸の値としたレーダーチャートを表示する。2 商品に共通する副評価項目が 3 つ未満である場合は副評価項目をラベル、その評判値を値とする棒グラフを表示する。また、副評価項目の中から 1 つの評価項目を選択することで、主評価項目を軸とするレーダーチャートの軸を選択した 1 つの副評価項目に置き換える。

また、図 1 上部の独自評判情報表示部に、製品のメーカー名と商品名、商品の画像と独自評価欄を表示する。独自評価欄には商品独自の評価項目を表示する。独自評価欄の上にマウスポインタを移動させると、評価項目ごとの評判値を Score として表示する。また、共通評判情報表示部において主評価項目を選択したとき、選択した主評価項目に関連する副評価項目の内、商品独自の副評価項目とその Score がそれぞれ商品の独自評価欄に表示される。

3.3 評価項目と評判値の生成

商品レビューから評価項目とその評判値を生成するために Google Natural Language API の analyzeEntitySentiment⁴メソッドを用いる。これは入力したテキストから感情極性が示されて

3 : <https://cloud.google.com/natural-language?hl=ja>

4 : <https://cloud.google.com/natural-language/docs/reference/rest/v1/documents/analyzeEntitySentiment>

2 : https://github.com/python/cpython/blob/3.11/Lib/tkinter/_init_.py

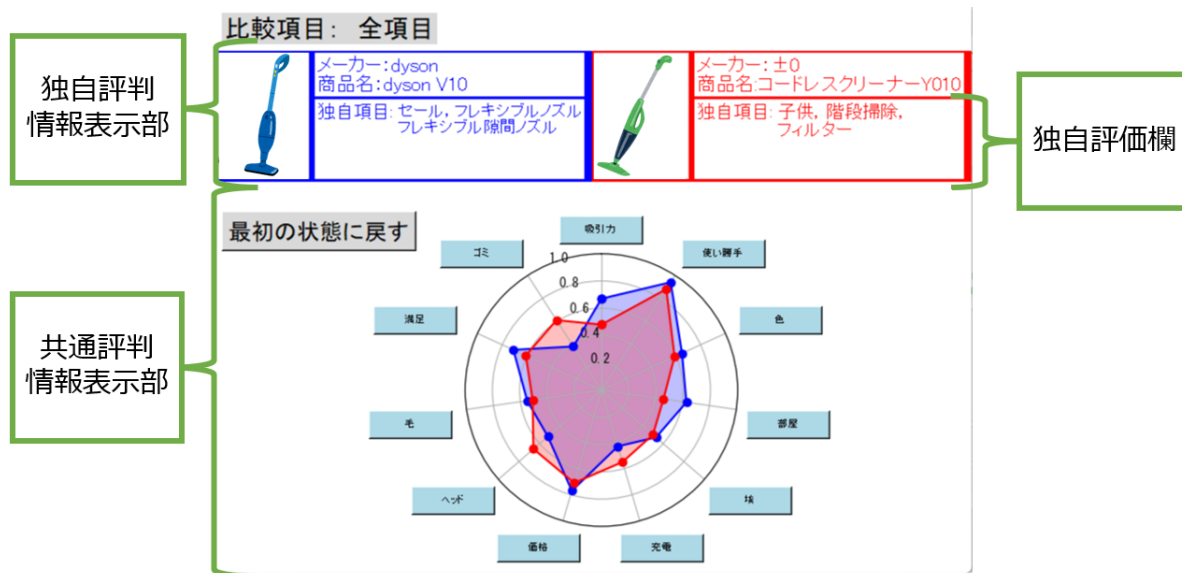


図1 提案手法のユーザインターフェースのイメージ

いる単語を Entity, その感情極性を数値化したものを Score, Entity の文章中の重要度を Saliency として出力する。Score は -1 から 1 までの実数として返され, Score は 1 に近いほどポジティブな感情であることを示し, -1 に近いほどネガティブな感情であることを示している。Saliency は 0 から 1 までの実数として返され, Saliency は 1 に近いほど Entity が文章において重要であることを示している。Google Natural Language API には比較したい 2 商品に投稿されたレビューすべてをまとめて渡す。得られた Entity の単語分散表現を Word2vec により獲得し, それをクラスタリングして評価項目を生成する。また, Score は評判値の生成に用いる。

3.3.1 不要語等の除去

Google Natural Language API から返される Entity には商品の評価項目として適切でない単語が存在する。また, 文章において一般語のように頻出する単語も多数存在する。そのため, そのような単語をストップワードとして除去する。除去する語は以下の語である。

- 「発送」や「梱包」などの配送に関する語 13 語
- 「商品レビュー」や「星 5」などの商品レビューそのものに関する語 11 語
- 「掃除機」, 「ビール」などの比較したい商品の商品ジャンルを示す語
- SlothLib ライブラリ [4] が公開している不要語 310 語

3.3.2 Word2Vec

本稿では単語の分散表現の獲得に Word2Vec を用いる。Word2Vec は Tomas ら [5] によって提案されたニューラルネットワークを用いた単語の分散表現の計算手法である。

本稿では Word2Vec に事前学習済みモデルである chiVe[6] を用いる。chiVe は約 1 億のウェブページ文章を含む国立国語研究所の日本語ウェブコーパス (NWJC) を学習しており, 分か

ち書きには形態素解析器 Sudachi[7] を用いている。

3.3.3 単語の分散表現のクラスタリング

Word2Vec によって得られた単語の分散表現を Ward 法によって階層的にクラスタリングする。Ward 法による階層型クラスタリングでは, まず, 各データを初期クラスタとする。結合によって生成されるクラスタの重心とクラスタ内の各データとの距離の二乗和と, 結合前の 2 つのクラスタ内での重心とそれぞれのデータとの距離の二乗和の差を比較し, その差が最小となるクラスタ同士を結合する。それを反復して実行し, 最終的に一つのクラスタになるまで繰り返す。階層型クラスタリングでは最小で 1 つのクラスタ, 最大でデータの総数のクラスタを得ることができ, クラスタを結合する基準とその閾値を設定するとその閾値以上のクラスタの結合を止め, 任意のクラスタ数にすることができる。

本研究では, クラスタ結合の基準を不整合度とする。2 つのクラスタの結合をリンク k , リンク k の高さを $Height(k)$, リンク k の下にあるすべてのリンクの高さの平均を $Average(k)$, リンク k の下にあるすべてのリンクの高さの標準偏差を $Standard Deviaton(k)$ とすると, リンク k の不整合度 $Inconsistence(k)$ は式 (1) で求められる。

$$Inconsistence(k) = \frac{Height(k) - Average(k)}{Standard Deviation(k)} \quad (1)$$

クラスタ結合の基準とする閾値はクラスタ内のリンクの不整合度の最大値 $\times 0.99$ とし, クラスタのリンクが閾値以上の不整合度を持つものは結合しない。

3.3.4 評価項目の生成

本研究では階層型クラスタリングの結果からそのクラスタを代表する単語を抽出し, 商品の主評価項目とする。主評価項目生成のためにまず, 階層型クラスタリングの初期クラスタの特徴を表すクラスタラベルをクラスタに含まれるデータである単語とする。次に, クラスタの結合の際にクラスタラベルを比較

し、片方だけが2つの商品のレビューに共通して出現するクラスタラベルであれば、そのクラスタラベルを結合した後のクラスタラベルとする。そうでなければ、Salience とクラスタラベルとなった単語のレビュー中の出現回数により定義した代表度をクラスタラベルごとに計算する。そして、その結果を比較し代表度が高いクラスタラベルを結合後のクラスタラベルとする。これを閾値によって決められたクラスタの最後の結合まで繰り返す。そのため、クラスタリングが終わったときのクラスタラベルはクラスタ内で最も代表度の高い単語となる。そして、そのクラスタラベルをその商品の主評価項目とする。また、商品と比較する上で評価項目が多すぎると比較が難しくなる。そのため、生成されたクラスタの内、クラスタ内のデータ数が2以下のものは評価項目としない。

主評価項目に関連する評価項目である副評価項目は以下のように生成する。階層型クラスタリングによって生成したクラスタごとにクラスタラベルを要素とするリストを生成する。そして、そのリストの先頭から8要素を副評価項目とする。リストの生成の例は図2のようになる。階層型クラスタリングによって生成されたクラスタ*i*内で一番最初に生成されたクラスタ*i-1*についての図下部のようなリストを生成する。結合されクラスタ*i-1*となった2つの初期クラスタのクラスタラベルである単語の内代表度が高いクラスタラベルをリストの先頭に挿入する。次に生成されたクラスタ*i-2*に含まれる2つの初期クラスタのクラスタラベルも代表度が高い方を前にしてリストの先頭に挿入する。最後に生成されたクラスタ*i*でも同様にクラスタ*i-1*とクラスタ*i-2*のクラスタラベルを比較し、代表度が高い方を前にしてリストの先頭に挿入する。このとき、リスト内に挿入したクラスタラベルと同じクラスタラベルをリストから取り除く。そして、生成したリストから図2の例の副評価項目は「吸引力」、「持ち」、「問題」、「音」となる。

なお、クラスタラベルの代表度は以下のように定義する。まず、商品レビュー中のクラスタラベルとなった単語の出現回数から出現回数が最小の単語の出現回数を引き、それを出現回数が最大の単語の出現回数から出現回数が最小の単語の出現回数を引いたもので割ることでクラスタラベルの出現回数を正規化させる。そして、その値とクラスタラベルのSalienceとの平均をそのクラスタラベルの代表度とする。これは0から1で表され、1に近いほどクラスタラベルになりやすいことを示す。

3.3.5 評価項目の評判値の生成

評価項目を生成した後、それぞれのクラスタ内の単語のScoreの平均を求める。そして、その値からScoreの最小値である-1を引き、Scoreの最大値である1からScoreの最小値である-1を引いた値で割ることでScoreの平均を正規化する。そして、その値を主評価項目の評判値とする。図1のユーザーインターフェースでは主評価項目を軸、算出した評判値を軸の値としたレーダーチャートを表示する。

また、副評価項目の評判値は副評価項目となっている単語のScoreも上のScore平均と同様に正規化し、評判値とする。

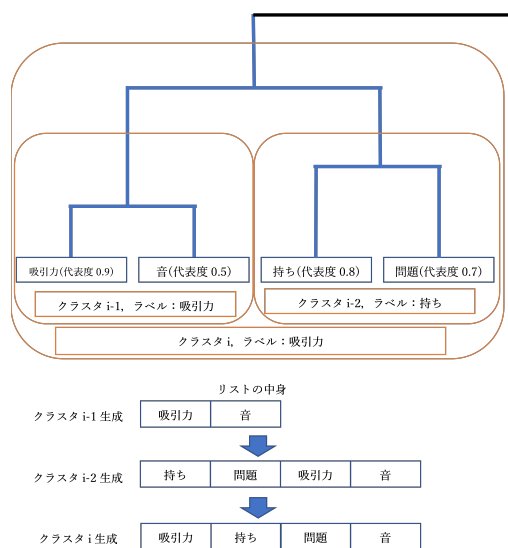


図2 クラスタ*i*における副評価項目のリストの生成

4 評価実験

4.1 Word2Vecのファインチューニング

本章の実験では、楽天グループ株式会社が提供している楽天市場データセット内の「みんなのレビュー・口コミ情報」[8]を使用して、Word2Vecの事前学習モデルであるchiVeをファインチューニングする。ファインチューニングには、2018年と2019年の間に楽天市場に投稿された商品ジャンル「掃除機」の45,005件のレビューと商品ジャンル「ビール」の23,173件のレビューを形態素解析器janome⁵によって分かち書きしたデータを用いる。

4.2 実験I: 主評価項目の有用性に関する実験

4.2.1 実験概要

実験Iでは、生成した主評価項目が評価項目として有用であるかを調べる。そのために、岡山大学工学部と岡山大学自然科学研究科の合わせて10人の学生に、提案手法により生成された主評価項目の数や内容の妥当性を評価させた。

実験では、被験者に提案手法によって生成された表1の「dyson V10」の主評価項目と表2の「アサヒ スーパードライ」の主評価項目の数の妥当性、重要さ、曖昧さを評価させる。そのアンケートの内容と回答の形式は表3にまとめる。また、質問の回答形式について、選択式では生成された主評価項目の中から項目を選択させる。5段階評価は、1から5までの整数を選択させる。この数値は1が「まったくそう思わない」、5が「とてもそう思う」である。また、記述式では質問の回答について自由に記述させる。

5: <https://github.com/mocobeta/janome>

表1 提案手法で生成した「dyson V10」の主評価項目

No.	主評価項目	評価項目に関連する単語
1	吸引力	吸引力, 持ち, 音, 問題
2	ゴミ	ゴミ, 部分, 本体, フタ
3	満足	満足, コードレス, コード, 結果, 稼働時間, V8
4	毛	毛, 髪
5	ヘッド	ヘッド, 壁
6	価格	価格, 満足度
7	充電	充電, バッテリー, 充電器
8	埃	埃, 汚れ
9	部屋	部屋, デザイン, スリム
10	色	色, 黒
11	使い勝手	使い勝手
12	セール	セール, 国産
13	フレキシブルノズル	フレキシブルノズル
14	フレキシブル隙間ノズル	フレキシブル隙間ノズル, スタンドクリーナー, 本体価格, 室内清掃

表2 提案手法で生成した「アサヒ スーパードライ」の主評価項目

No.	主評価項目	評価項目に関連する単語
1	味	味
2	値段	値段, 価格, ショップ
3	物	物, 段ボール, 搾り, 誕生日プレゼント
4	海外ビール	海外ビール
5	クーポン	クーポン, 定番, ストック, 満足度
6	缶	缶, 箱
7	おまけ	おまけ, オフクーポン

表3 提案手法で生成した主評価項目に関するアンケート

No.	質問内容	回答方式
1	評価項目は多いと感じたか.	5段階評価
2	どの評価項目を必要ないと思うか.	選択式
3	評価項目は少ないと感じたか.	5段階評価
4	追加してほしい評価項目はどんな評価項目か.	記述式
5	どの評価項目が重要だと感じたか.	選択式
6	どの評価項目の内容が曖昧だと感じたか.	選択式
7	より詳しく知りたいと感じた評価項目をすべて選択せよ.	選択式
8	7で選択した評価項目のどのような点を詳しく知りたいか.	記述式

4.2.2 実験結果

表4に2つの商品について、評価項目の数の妥当性を示す。表5には、「dyson V10」で生成された評価項目ごとに表3の質問でその評価項目を選択した被験者の人数をまとめる。また、表6には、「アサヒ スーパードライ」の評価項目ごとに表3の質問でその評価項目を選択した被験者の人数をまとめる。

表4を見ると、提案手法によって生成した「dyson V10」の主評価項目の数は多いと評価されている。また、必要ない評価項目として「ゴミ」、「毛」、「ヘッド」、「部屋」、「フレキシブルノズル」、「フレキシブル隙間ノズル」を選択した被験者が多かった。これらの項目の内「ゴミ」、「毛」、「部屋」、「フレキシブル隙間ノズル」は曖昧な評価項目とした被験者も多かった。これらの評価項目には関連する単語として関連が不明なものも多く集められていた。また、追加したい項目として「音」や「重量」が複数の被験者から挙げられた。また、「吸引力」、「価格」、「充電」、「使い勝手」を重要な評価項目とした被験者が多かった。これらの評価項目は多くの被験者が詳しく知りたい評価項目とし、特に「吸引力」、「使い勝手」に関しては具体的にどの吸引力があるのかや、「ヘッドの使い勝手」など具体的にどの部分

表4 評価項目の数に関する質問への回答

質問	dyson V10	アサヒ スーパードライ
評価項目は多いと感じたか.	3.7	2.9
評価項目は少ないと感じたか.	1.4	1.8

表5 提案手法で生成した「dyson V10」の主評価項目と各質問における選択人数

主評価項目	表3の質問番号			
	2(不要)	5(重要)	6(曖昧)	7(詳細が知りたい)
吸引力	0	10	1	4
ゴミ	4	2	7	2
満足	1	3	5	3
毛	6	1	6	1
ヘッド	5	0	4	1
価格	0	10	0	3
充電	0	8	0	4
埃	4	3	5	1
部屋	8	1	7	1
色	3	1	1	1
使い勝手	0	7	4	5
セール	3	2	5	2
フレキシブルノズル	5	0	4	1
フレキシブル隙間ノズル	5	0	5	2

の使い勝手がいいのを知りたいといった意見が寄せられた。

一方、表4では、提案手法によって生成した「アサヒ スーパードライ」の評価項目の数はおおむね適切であると評価されている。また、必要ない項目としては「物」、「缶」、「おまけ」を選択する被験者が多かった。これらの項目は曖昧な項目としても多く選択された。重要な項目としては「味」、「値段」が多く選択され、特に「味」に関しては詳しく知りたい項目であるとする被験者も多かった。具体的に知りたい内容としては、のどごしやキレ、苦みや甘みなど味に関する具体的な評価が挙げられた。

4.2.3 考察

商品ジャンル「掃除機」の評価項目数は多いと評価された。それと比較し商品ジャンル「ビール」の評価項目数は適切と評価された。商品ジャンル「掃除機」において評価項目の数を減らすためには、階層型クラスタリングの適切な閾値を設定することなどが必要である。

また、どちらの商品でも重要だと評価される評価項目はあったが、その数は評価項目の総数の半数に満たなかった。この一因は、商品の評価項目として曖昧と思われる評価項目を多く生成したためである。加えて、商品ジャンル「掃除機」では追加してほしい評価項目として被験者の複数人から「音」と「重量」が挙げられた。このことから、生成した評価項目には必要な評価項目が不足していることも分かった。

4.3 実験 II: 提案手法を用いた商品比較に関する実験

4.3.1 実験概要

実験 II では、提案手法が商品の比較において役立つかどうかを評価するため、実験 I と同じ被験者 10 人に、価格.com のインターフェースと提案手法を比較させる。そして、被験者が商

表6 提案手法で生成した「アサヒ スーパードライ」の評価項目と各質問における選択人数

主評価項目	表3の質問番号			
	2(不要)	5(重要)	6(曖昧)	7(詳細が知りたい)
味	0	10	3	8
値段	0	9	0	2
物	7	0	9	0
海外ビール	3	0	4	0
クーポン	3	1	4	0
缶	5	0	5	0
おまけ	7	1	5	1

表7 実験IIで被験者の各グループが実施する内容の順番

実施順	グループ A	グループ B
1	価格.comで「掃除機」の2商品比較	価格.comで「ビール」の2商品比較
2	商品の比較に役立った項目選択	
3	提案手法で「掃除機」の2商品比較	提案手法で「ビール」の2商品比較
4	商品の比較に役立った項目選択	
5	比較しやすかった手法選択	
6	提案手法で「ビール」の2商品比較	提案手法で「掃除機」の2商品比較
7	商品の比較に役立った項目選択	
8	価格.comで「ビール」の2商品比較	価格.comで「掃除機」の2商品比較
9	商品の比較に役立った項目選択	
10	比較しやすかった手法選択	

品を比較し終えたら、各手法について商品と比較する上でどの評価項目が役に立ったかとどちらの比較手法が使いやすかったかを尋ねる。

実験ではまず、被験者を5人ずつの2つのグループに分ける。これは各手法の利用の順番による質問の回答への影響を考慮するためである。そして、被験者に商品ジャンル「掃除機」の「dyson V10」と「±0コードレスクリーナー Y010」、「ビール」の「アサヒ スーパードライ」と「ヒューガルデン ホワイト」を提案手法と価格.comの両方を使って比較させる。その後、商品の比較に役立った評価項目をすべて選択させ、商品と比較しやすかった手法として「提案手法」、「どちらでもない」、「価格.com」の3つの内1つを選択させる。価格.comには提案手法での評価項目と同様に他のユーザが項目ごとに商品进行评估した評価点数という種類の評価項目があり、その他に提案手法にはない商品の具体的なスペックを表す評価項目もある。そこで、被験者に価格.comと提案手法を比較させる際、条件を揃えるために評価点数のみを評価項目とした場合の価格.comと提案手法を比較させる。価格.comで種類が評価点数である評価項目は「掃除機」では「満足度」、「デザイン」、「使いやすさ」、「吸引力・パワー」、「静音性」、「サイズ」、「手入れのしやすさ」、「取り回し」の8つであり、「ビール」では「満足度」、「キレ」、「コク」、「苦味」の3つである。被験者の各グループが実験で行う内容を実施順に表7にまとめる。

価格.comで商品ジャンル「掃除機」と「ビール」の2商品と比較した後に選択させる評価項目の選択肢を表8に示す。また、提案手法で商品ジャンル「掃除機」の2商品と比較した後に選択させる評価項目の選択肢を表9に、商品ジャンル「ビール」の2商品と比較した後のそれらを表10に示す。

表8 実験IIで被験者に選択させる価格.comの「掃除機」と「ビール」評価項目

商品ジャンル	評価項目(選択肢)
掃除機	製品名、メーカー、製品画像、色、最安値、ランキング、発売日、クチコミ件数、満足度、デザイン、使いやすさ、吸引力・パワー、静音性、サイズ、手入れのしやすさ、取り回し、満足度4以上、満足度2以下、価格推移グラフ、集じん方式、サイクロン種類、集じん容積、タイプ、2in1(2way)、コードレス、ヘッド種類、最長運転時間(連続使用時間)、充電時間、ダストケース丸洗い、本体寸法(幅×高さ×奥行)、質量
ビール	製品名、メーカー、製品画像、最安値、ランキング、発売日、クチコミ件数、満足度、キレ、コク、苦味、満足度4以上、満足度2以下、価格推移グラフ、種類、容器、容量、本数、1本当たりの価格、アルコール分、エネルギー、産地

表9 実験IIで被験者に選択させる提案手法で生成した「掃除機」の評価項目(総数79)

評価項目(選択肢)	評価項目の種類	
	共通または独自(商品名)	主評価項目または副評価項目(主評価項目)
吸引力、ゴミ、満足、ヘッド、価格	共通	主評価項目
充電、埃、部屋、色、使い勝手	共通	
セール、フレキシブルノズル、フレキシブル隙間ノズル	dyson V10	副評価項目(吸引力)
子供、階段掃除	±0 Y010	
吸引力、持ち、音、問題	共通	副評価項目(吸引力)
レベル	±0 Y010	
ゴミ、部分、本体	共通	副評価項目(ゴミ)
フタ、ダストボックス、ダストカップ	±0 Y010	
満足、コードレス、コード	共通	副評価項目(満足)
結果、稼働時間、V8	dyson V10	
使用、紙パック式	±0 Y010	副評価項目(毛)
毛、髪	共通	
カーペット、階段、抜け毛、フローリング、廊下、絨毯	±0 Y010	副評価項目(ヘッド)
ヘッド	共通	
壁	dyson V10	副評価項目(ヘッド)
移動	±0 Y010	
価格	共通	副評価項目(価格)
満足度	dyson V10	
コスト、値段	±0 Y010	副評価項目(充電)
充電、バッテリー、充電器	共通	
電池パック、電池	±0 Y010	副評価項目(埃)
埃	共通	
汚れ	dyson V10	副評価項目(埃)
クリーナー	±0 Y010	
部屋、デザイン	共通	副評価項目(部屋)
スリム	dyson V10	
スタンド、インテリア、フォルム、リビング、スタイリッシュ	±0 Y010	副評価項目(色)
色	共通	
黒	dyson V10	副評価項目(色)
実物、グレー、真っ白、カラー、ベージュ、ピンク色	±0 Y010	
使い勝手	共通	副評価項目(使い勝手)
使い心地、勝手	±0 Y010	

4.3.2 実験結果

価格.comでの「掃除機」の商品比較に役立った評価項目とそれを選択した人数を表11に示す。「製品画像」、「最安値」などの基本的な情報を示す評価項目を比較に役立った評価項目とした被験者が多かった。他には他者からの商品に対する評価が関連する「ランキング」や「クチコミ件数」、「満足度4以上」、「満足度2以下」の評価項目も同様に役立った評価項目とした被験者が多かった。

表 10 実験 II で被験者に選択させる提案手法で生成した「ビール」の評価項目 (総数 36)

評価項目 (選択肢)	評価項目の種類	
	共通または独自 (商品名)	主評価項目または副評価項目 (主評価項目)
味, 値段, 物, 海外ビール	共通	主評価項目
香り, 賞味期限, グラス, 瓶入り, 癖, 品質, 後味, 爽やか白ビール, コスパ	ヒューガルデン ホワイト	
クーポン, 缶, おまけ	アサヒ スーパードライ	
味	共通	副評価項目 (味)
美味, 格安, 味わい, 風味	ヒューガルデン ホワイト	副評価項目 (値段)
値段, 価格	ヒューガルデン ホワイト	
販売	ヒューガルデン ホワイト	副評価項目 (物)
ショップ	アサヒ スーパードライ	
物, 段ボール	共通	副評価項目 (海外ビール)
搾り, 誕生日プレゼント	アサヒ スーパードライ	
海外ビール	共通	副評価項目 (海外ビール)
苦味, ホワイトビール, 専用グラス, 白ビール, 輸入ビール, フルーティーさ	ヒューガルデン ホワイト	

一方, 提案手法で生成した「掃除機」の評価項目を役立った項目と選択した人数ごとに表 12 に示す。10 人の内 5 人以上が選択した評価項目は評価項目数の半数以下だった。また, 役立った主評価項目の副評価項目を役立った項目とした被験者が比較的多かった。しかし, 役に立たなかったとされた主評価項目の「部屋」や「色」の商品独自の副評価項目の中に 5 人以上が役立つとした副評価項目がいくつかあった。

価格.com で「ビール」の商品比較に役立った評価項目とそれを選択した人数を表 13 に示す。「最安値」, 「1 本当たりの価格」といった値段に関する評価項目が役立ったとされた。「満足度」は役立った項目として 10 人全員が選択した。しかし, 「キレ」, 「コク」, 「苦味」の味に関する具体的な評価項目は半数ほどしか選択されなかった。

一方, 提案手法で生成した「ビール」の評価項目を役立った項目と選択した人数ごとに表 14 に示す。10 人の内 5 人以上が選択した評価項目は「掃除機」と同様に評価項目数の半数以下だった。「味」と「値段」を比較に役立った主評価項目とした被験者が多かった。また, ほとんどの副評価項目は役立たないとされた。

2 つの手法による比較を行った上での質問「価格.com と提案手法を比べてどちらの比較手法が比較しやすかったか。」に対する回答を表 15 に示す。全体的に価格.com を比較しやすいとした被験者が多かったが, 「掃除機」では 10 人の内 2 人が提案手法を選択した。

4.3.3 考察

実験 II の結果から, どちらの商品ジャンルでも価格.com より提案手法が商品と比較しづらかったことが分かる。その理由の 1 つとして, 比較に役立った項目数が少ないことが挙げられる。提案手法が比較する上で不要な評価項目も多く生成したことが一因であるため, 不要な評価項目を抑制する必要がある。

また, 「ビール」と違い, 「掃除機」では一部の被験者が提案手法を支持した。提案手法において「掃除機」では商品独自の副評価項目が役立った評価項目とされた。これらの商品独自の

表 11 価格.com における「掃除機」の評価項目とそれを役立った評価項目と選択した人数

選択人数	評価項目
0	—
1	—
2	デザイン, 集じん方式
3	サイズ, サイクロン種類, 集じん容積
4	ダストケース丸洗い
5	製品名, 色, 発売日, 取り回し, 本体寸法 (幅×高さ×奥行)
6	使いやすさ, タイプ
7	メーカー, クチコミ件数, 静音性, 手入れのしやすさ, 最長運転時間/連続使用時間
8	吸引力・パワー, 満足度 4 以上, 満足度 2 以下, コードレス, 充電時間, 質量
9	製品画像, ランキング, 満足度
10	最安値

表 12 提案手法で生成した「掃除機」の評価項目とそれを役立った評価項目と選択した人数

選択人数	評価項目
0	子供, 問題, レベル, 部分, 結果, V8, リビング, 実物
1	毛 (主評価項目), 本体, フタ, 使用, 毛 (副評価項目), 抜け毛, ヘッド (副評価項目), 壁, フォルム (副評価項目), 価格 (副評価項目), 電池, クリーナー, 部屋 (副評価項目),
2	ヘッド (主評価項目), 埃 (主評価項目), 部屋 (主評価項目), 色 (主評価項目), 持ち, ゴミ (副評価項目), 移動, 充電器, 電池パック, 埃 (副評価項目), 汚れ, インテリア, カラー
3	ゴミ (主評価項目), セール, フレキシブルノズル, フレキシブル隙間ノズル, フィルター, ダストカップ, コード, 紙パック式, 髪, カーペット, フローリング, 廊下, 絨毯, スタンド, 色 (副評価項目)
4	階段掃除, ダストボックス, 満足 (副評価項目), コードレス, 階段, 真っ白, 使い心地
5	稼働時間, 満足度, 充電, スリム, スタイリッシュ, 黒, グレー, ベージュ, ピンク色, 使い勝手 (副評価項目)
6	価格 (主評価項目), 充電 (主評価項目), コスパ, 値段, バッテリー, デザイン,
7	吸引力 (主評価項目), 満足 (主評価項目), 使い勝手 (主評価項目), 音
8	吸引力 (副評価項目)
9	—
10	—

副評価項目は役に立たないとされた主評価項目に関連する評価項目が多かった。そのうち価格.com では役立った項目とされなかった「デザイン」に関する商品独自の副評価項目も役立った評価項目とされた。価格.com ではデザインに関する評価項目は「デザイン」のみとなっており, 提案手法では「スリム」や「スタイリッシュ」といった「デザイン」のより具体的な評価項目の評判も可視化している。このことから提案手法では「掃除機」のジャンルにおいて商品比較に役立った商品独自の副評価項目を生成したといえる。

5 まとめ

本稿では, ユーザが 2 商品と比較する状況を想定し, 商品レビューから商品ジャンルごとに自動生成した評価項目の評判を, レーダーチャート等で可視化する手法を提案した。提案手法では, レビューから抽出した単語を階層型クラスタリングして, 2 商品に共通する主評価項目と各商品独自の主評価項目を生成し, 主評価項目に関連する副評価項目も生成した。また提

表 13 価格.com における「ビール」の評価項目とそれを役立った評価項目と選択した人数

選択人数	評価項目
0	—
1	発売日
2	エネルギー
3	種類
4	ランキング, 価格推移グラフ
5	キレ, 苦味, 容器, 産地
6	製品名, メーカー, クチコミ件数, コク, アルコール分
7	満足度 2 以下, 容量, 本数
8	—
9	満足度 4 以下, 1 本当たりの価格
10	製品画像, 最安値, 満足度

表 14 提案手法で生成した「ビール」の評価項目とそれを役立った評価項目と選択した人数

選択人数	評価項目
0	販売, ショップ, 段ボール, 搾り, 海外ビール (主評価項目)
1	物 (主評価項目), 海外ビール (主評価項目), グラス, クーポン, おまけ, 物 (副評価項目), 誕生日プレゼント
2	賞味期限, 味わい, 価格, 輸入ビール, フルーティさ
3	品質, 爽やか白ビール, 缶, 美味, 風味, 値段 (副評価項目), 苦味, ホワイトビール, 専用グラス, 白ビール
4	瓶入り, 後味, 格安
5	癖, コスパ, 味 (副評価項目)
6	—
7	値段 (主評価項目), 香り
8	—
9	味 (主評価項目)
10	—

表 15 価格.com と提案手法を比べてどちらの比較手法が比較しやすかったか.

商品ジャンル	提案手法	どちらでもない	価格.com
掃除機	2	2	6
ビール	0	1	9

案する商品レビューの可視化インターフェースでは、ユーザは副評価項目を主評価項目と入れ替えることで自分が比較したい評価項目で商品を比較できる。

実験では、提案手法により生成した評価項目の有用性と商品比較の場面での有用性を検証するために、大学生と大学院生あわせて 10 名による被験者実験を実施した。被験者は提案手法によって生成された評価項目の数と内容の妥当性を評価した。結果として商品ジャンル「掃除機」では、生成された評価項目の数が多く、重要な項目が生成されたが曖昧な項目も多く生成された。よって、今後の課題として商品レビューから抽出した単語のクラスタリングの精度を上げ、適切なクラスタを生成することが挙げられる。実験ではまた、被験者に提案手法と既存の商品比較サイト価格.com で同じ 2 商品を比較させ、それぞれ

の手法で比較に役立った要素とどちらの手法が比較しやすかったかを問うた。その結果、価格.com の方が商品比較がしやすいと感じる被験者が多かった。しかし、提案手法は商品ジャンル「掃除機」の 2 商品の比較において、被験者が有用と感じるデザインに関する評価項目を生成できた。

謝 辞

本稿では国立情報学研究所の IDR データセット提供サービスにより楽天グループ株式会社から提供された「楽天データセット」(https://rit.rakuten.com/data_release/) を利用した。ここに記して感謝する。

文 献

- [1] 藤本大吾, 上野史, 太田学, “評判情報を用いた代替品推薦の一手法,” DEIM2022, C34-2, 2022.
- [2] 平山拓史, 湯本高行, 新居学, 佐藤邦弘, “語の共起と極性に基づく商品レビュー閲覧支援システム,” 情報処理学会研究報告, Vol. 2012-DBS-155, No. 3, pp. 1-9, 2012.
- [3] 矢部沙也加, 上田真由美, 中島伸介, “評価項目別スコアを用いたコスメアイテム間の相違点可視化システム,” DEIM2021, C14-2, 2021.
- [4] 大島 裕明, 中村 聡史, 田中 克己, “SlothLib: Web サーチャ研究のためのプログラミングライブラリ,” 日本データベース学会 Letters, Vol. 6, No. 1, pp. 113–116, 2007.
- [5] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean, “Distributed Representations of Words and Phrases and their Compositionality,” In Proceedings of the NIPS 2013, Vol. 2, pp. 3111–3119, 2013.
- [6] 真鍋陽俊, 岡照晃, 海川祥毅, 高岡一馬, 内田佳孝, 浅原正幸, “複数粒度の分割結果に基づく日本語単語分散表現,” 言語処理学会第 25 回年次大会, NLP2019-P8-5, 2019.
- [7] Kazuma Taoka, Sorami Hisamoto, Noriko Kawahara, Miho Sakamoto, Yoshitaka Uchida, Yuji Matsumoto, “Sudachi: a Japanese Tokenizer for Business,” Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 2018.
- [8] 楽天グループ株式会社, “楽天市場データ. 国立情報学研究所情報学研究データリポジトリ (データセット),” <https://doi.org/10.32130/idr.2.1> (2022 年 6 月 19 日閲覧).