

## 修士論文の和文要旨

研究科・専攻	大学院 情報理工学研究科 情報・通信工学専攻 博士前期課程		
氏名	金兵裕太	学籍番号	1531026
論文題目	ネットショッピングサイトの商品レビューを利用した ジャンル毎の評価軸の自動構築とその応用		
要旨	<p>SNS やブログに代表される、消費者が形成するメディアを <b>CGM (Consumer Generated Media)</b> と呼ぶ。インターネットが普及した現代においては、<b>CGM</b> は消費者と企業双方にとって有益なものであり、この <b>CGM</b> を分析する動きは盛んになっている。</p> <p>代表的な <b>CGM</b> の一つであるネットショッピングサイトの商品レビューは、商品の購入を決定する際の手助けとなるものであるが、大量のレビューが存在する場合、その全てを読むことは困難である。そのため、自動的に商品レビューを分析し、共通の判断基準に基づいて要約する試みは数多くある。商品レビューを始めとする評価文書が示す、肯定や否定の感情を分析する分野は "評判分析"や"感情分析"と呼ばれ、要約や情報推薦を支える技術である。<b>CGM</b> の増大に伴い、このような情報推薦の技術にも注目が集まっており、効果的な情報推薦への需要が高まっている。</p> <p>しかし、既存の評価表現辞書を用いた評判分析手法では、レビューを肯定的か否定的かでしか評価しないため、商品に存在する複数の評価指標を反映することが出来ない。そのため、ユーザに提供できる情報量が少ないという問題がある。評価対象となる特徴語や属性を判定する研究も存在するが、評価対象をカテゴリ化するには予めドメイン毎の種語集合が必要であった。</p> <p>そこで本論文では、商品レビュー集合をコーパスとした評価軸と評価表現辞書の自動構築手法、及び分析手法を提案する。商品レビューから学習を行うことにより、「掃除機」や「プリンタ」といった商品ジャンルに特化した評価軸情報を構築することを可能とした。また、それを利用した評判分析を行うことで多面的な評価を可能にし、レーダーチャートの自動生成を実現した。</p>		

# 平成 28 年度修士論文

## ネットショッピングサイトの商品レビューを利用した ジャンル毎の評価軸の自動構築とその応用

電気通信大学大学院  
情報理工学研究科 情報・通信工学専攻  
コンピュータサイエンスコース

学籍番号 : 1531026

氏名 : 金兵裕太

指導教員 : 沼尾雅之 教授

副指導教員 : 岩崎英哉 教授

提出日 : 2017 年 1 月 30 日

## 要旨

SNS やブログに代表される、消費者が形成するメディアを CGM (Consumer Generated Media) と呼ぶ。インターネットが普及した現代においては、CGM は消費者と企業双方にとって有益なものであり、この CGM を分析する動きは盛んになっている。

代表的な CGM の一つであるネットショッピングサイトの商品レビューは、商品の購入を決定する際の手助けとなるものであるが、大量のレビューが存在する場合、その全てを読むことは困難である。そのため、自動的に商品レビューを分析し、共通の判断基準に基づいて要約する試みは数多くある。商品レビューを始めとする評価文書が示す、肯定や否定の感情を分析する分野は"評判分析"や"感情分析"と呼ばれ、要約や情報推薦を支える技術である。CGM の増大に伴い、このような情報推薦の技術にも注目が集まっており、効果的な情報推薦への需要が高まっている。

しかし、既存の評価表現辞書を用いた評判分析手法では、レビューを肯定的か否定的かでしか評価しないため、商品に存在する複数の評価指標を反映することが出来ない。そのため、ユーザに提供できる情報量が少ないという問題がある。評価対象となる特徴語や属性を判定する研究も存在するが、評価対象をカテゴライズするには予めドメイン毎の種語集合が必要であった。

そこで本論文では、商品レビュー集合をコーパスとした評価軸と評価表現辞書の自動構築手法、及び分析手法を提案する。商品レビューから学習を行うことにより、「掃除機」や「プリンタ」といった商品ジャンルに特化した評価軸情報を構築することを可能とした。また、それを利用した評判分析を行うことで多面的な評価を可能にし、レーダーチャートの自動生成を実現した。

# 目次

第1章	はじめに	1
1.1	背景	1
1.2	研究の目的	1
1.3	評判分析システム	2
1.4	本論文の構成	3
第2章	関連研究	5
2.1	評価表現辞書の構築に関する研究	5
2.2	属性等の付加情報を抽出する研究	6
2.3	評判分析に関する研究	6
2.4	既存研究における問題点	7
第3章	用語の定義	8
3.1	評価表現辞書	8
3.2	評価軸	9
3.3	商品ジャンル	9
第4章	語彙情報の学習	10
4.1	学習手順	10
4.2	語彙の収集	11
4.2.1	評価表現候補の抽出	11
4.2.2	特徴語候補の抽出	12
4.2.3	候補の抽出例	13
4.2.4	評価表現候補の評価極性値の算出	14
4.2.5	特徴語候補の重要度の算出	14
4.2.6	判定条件によるフィルタリング	15
4.3	評価軸の構築	17
4.3.1	特徴語間の距離の定義	17
4.3.2	特徴語の分類	19
4.3.3	評価軸の後処理	20

第5章 レーダーチャート生成システム .....	21
5.1 システム構成.....	21
5.2 レーダーチャート生成の手順.....	22
5.3 各レビューの評価値の算出方法.....	22
第6章 評価実験 .....	26
6.1 実験条件.....	26
6.1.1 楽天データセット .....	26
6.1.2 閾値.....	27
6.2 評価表現辞書の構築.....	28
6.2.1 語彙の収集結果.....	28
6.2.2 評価極性値の評価 .....	30
6.2.3 語彙数の評価 .....	33
6.3 特徴語の収集結果.....	36
6.4 評価軸の構築.....	38
6.5 商品の評判分析.....	46
6.5.1 レーダーチャートによる可視化.....	46
第7章 おわりに .....	48
7.1 まとめ.....	48
7.2 今後の課題.....	49
謝辞.....	50
参考文献.....	51
付録 評判分析システム .....	53

# 第 1 章 はじめに

## 1.1 背景

SNS やブログに代表される、消費者が形成するメディアを CGM (Consumer Generated Media) と呼ぶ。CGM は、消費者の購買傾向や、ユーザの商品・サービスに対する評価といった情報を有しており、今や企業にとって欠かせないものである[1]。Web 上の CGM を分析することで、企業は施策の立案、実施、確認、改善を短いサイクルで回すことが出来るようになるため、あらゆる分野で CGM 分析が盛んになっている[2]。

一方で、この CGM は消費者にとっても有益なものである。代表的な CGM の一つであるネットショッピングサイトの商品レビューは、ユーザが商品の購入を決定する際の手助けとなるものである。インターネットが普及したことにより、不特定多数のユーザからの評価を容易に見ることが可能となったため、ユーザは従来では知り得なかった情報を事前に得ることが出来るようになった。

しかし、レビューが大量に存在する場合、その全てを読むことは困難である。また、商品レビューには評価の点数が付与されているものが多いが、点数はレビュアーが個人の判断で付けるため、レビュー内容を正しく反映しているとは限らない。そのため、自動的にレビューを分析し、共通の判断基準で要約、可視化することは、ユーザが商品を比較検討する際に有益である。

商品レビューを始めとする評価文書が示す、肯定や否定の感情を分析する分野は"評判分析"や"感情分析"と呼ばれる。CGM の増大に伴い、効果的な情報推薦は今後さらに重要になると考えられ、盛んに研究が行われている。評判分析に関する既存の研究には、肯定語と否定語を登録した評価表現辞書を用いる手法等が存在するが、これらは「レビューが肯定的か否定的か」という 1 次元の尺度でしか評価することができない。しかし実際には、商品には「性能」や「価格」といった複数の評価軸が存在するため、肯定と否定の情報だけでは不十分だと考えられる。

## 1.2 研究の目的

本研究の目的は、ユーザの商品選択を支援するシステムを作成することである。商品に寄せられたレビューを自動的に要約・可視化することで、ユーザに対し商品の評価情報を簡潔に提示する。

我々は、評価軸は「掃除機」や「プリンタ」といった商品のジャンルによって異なることに着目し、商品ジャンル毎のレビューから評価軸と評価表現辞書を自動構築する手法を提案する。また、構築した評価軸に基づき商品レビューを分析することで「性能」や「価格」といった項目ごとの評価を可能にした。これにより、商品ジャンル毎に重要な評価指標を判別し、レーダーチャートを自動生成することが可能になる。従って、ユーザは既存手法の 1 次元的な評価値ではなく、多次元的な評価値を比べることができるようになり、類似した商品同士の比較検討が容易になると期待できる。

### 1.3 評判分析システム

図 1.1 は、本研究で想定する商品選択支援システムの使用例である。プリンタの評価軸として「価格」や「印刷速度」といった項目を設定する事で類似商品同士の比較検討が容易になり、購入者は自分の価値観に合った商品を既存の分析システムよりもスムーズに選択する事ができる。

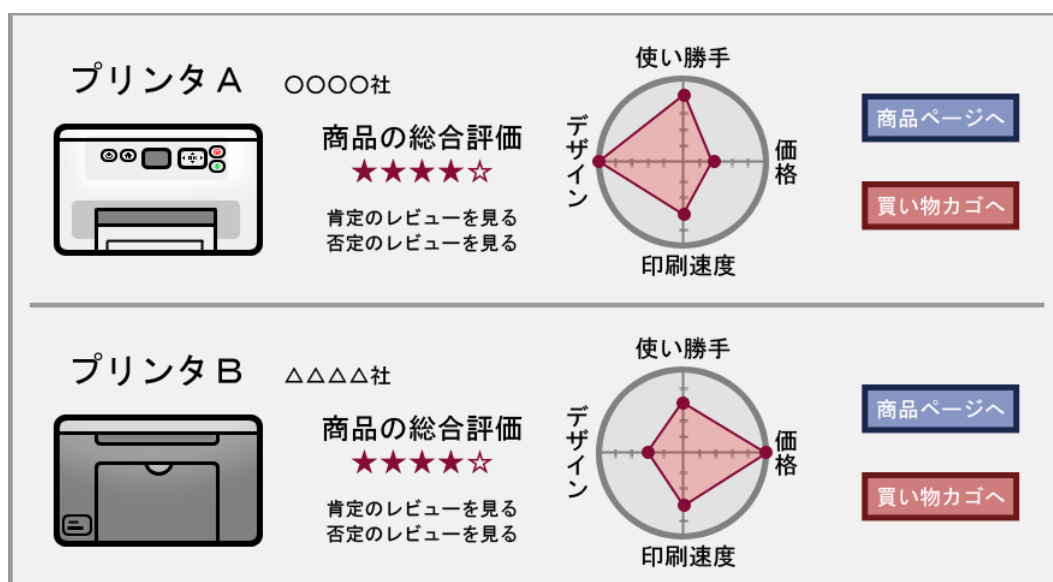


図 1.1 商品選択支援システムの例

本研究ではこのような多面的な評価を、評価表現辞書と特徴語を用いることで実現している。ここで特徴語とは、特定の評価指標を表現する語句である。レビュー中の特徴語が肯定の評価をされているか、否定の評価をされているかに応じて、特徴語の評価値を計算する。例として、「この価格は嬉しいです」という表現では、特徴語「価格」が評価表現「嬉しい」によって肯定の評価をされていると判定できる。しかし特徴語の中には「価格」と

「値段」のように、同様の評価指標を表現するものが多数存在する。そのような特徴語は個別に評価せず、同様の軸で扱うべきである。

そこで我々は、評価軸を「特徴語の集合」として定義する。図 1.2 のように、同様な意味を持つ特徴語同士を同じ評価軸として扱うことで、より簡便で正確な可視化を行うことができると考えられる。

このための主要な課題は、特定のジャンルの商品レビュー集合を入力として与えた時、そのジャンルに対応した評価表現辞書と特徴語を学習し、評価軸を構築することである。そして最終的には構築した語彙情報を用いて商品レビューを分析し、レーダーチャートを生成する。この実現にあたり、特徴語と評価表現の間の関係を抽出するために、係り受け解析を利用する。

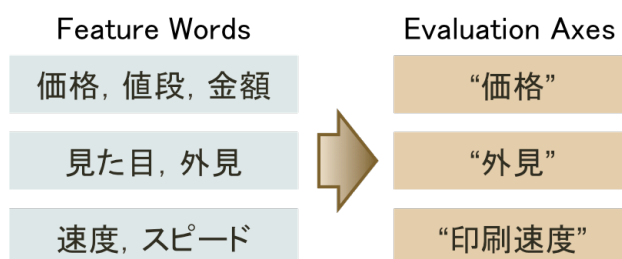


図 1.2 評価軸と特徴語の例

## 1.4 本論文の構成

### 第 1 章

本研究の背景と目的について説明する

### 第 2 章

関連研究について説明する

### 第 3 章

本研究で扱う用語について説明する

### 第 4 章

語彙の収集方法について説明する

### 第 5 章

収集した語彙を用いてレーダーチャートを生成する方法について説明する



## 第 6 章

提案手法の評価実験を行う

## 第 7 章

本研究のまとめを行う

## 第 2 章 関連研究

### 2.1 評価表現辞書の構築に関する研究

評判情報を分析するには、ユーザの評価が現れる評価表現を扱う場合が多く、その辞書を構築するための取り組みは多岐に渡る[3].

評価表現辞書を人手により構築する取り組みは複数存在する. 大学共同利用機関法人人間文化研究機構国立国語研究所コーパス開発センターが発表した「日本語アプレイザル評価表現辞書」は、評価表現に肯定や否定という情報を付与するだけでなく、語義に基づき複数のカテゴリに分類している. 表現が評価対象の特徴を表すものか、評価者の感情や行為を示すものかといった情報を得ることが出来るため、幅広い研究分野への活用が可能だとされている[4]. また、小林らは意見マイニングを目的として、分野横断的に利用可能な汎用的な評価表現辞書を公開している[5]. 意見タグつきコーパスを作成することで、半自動的に評価表現辞書を作成することが可能である. 意見タグつきコーパスは、性質の異なる 4 つのドメインを対象にして作成されており、作成した評価表現辞書がどのドメインにおいても表現の約 9 割をカバーできたと報告している. しかし、対象とした文書集合が比較的均質であった可能性等を示唆しており、どのドメインに対しても汎用的に利用可能かは疑問が残る.

これらの人手により構築した評価表現辞書は、分野横断的に利用するという前提があるため、公開された辞書を入手すればすぐに活用できるという大きなメリットがあるが、ドメインに依存する表現をカバーできない可能性を孕んでいるほか、人手による構築に膨大なコストがかかるという欠点がある. こういった側面から、評価表現辞書は自動的に構築する必要があるとする声も多く、そのための研究も数多く存在する. Kamps らは肯定語と否定語を収集する手法として、形容詞の類義語関係の語彙ネットワークを利用する手法を提案した[6]. 肯定と否定を代表する語として「good」と「bad」を選定し、判定したい形容詞が、語彙ネットワーク内で「good」と「bad」どちらに近いかを計算することで、肯定・否定の極性を付与するというものである. しかし、語彙ネットワーク内の言葉しか登録できないため、新語などの未知語には対応できないといった欠点があった. 一方、那須川らは周辺文脈の情報を利用した手法を提案している[7]. この手法では、評価表現の文脈一貫性という仮定に基づき、共起情報を用いて種表現からブートストラップ的に評価表現候補を収集する. 結果として、分野特有の語彙を高い精度で抽出することに成功している.

## 2.2 属性等の付加情報を抽出する研究

評価表現辞書の構築の延長線上に、評価表現と共に属性等の付加情報を抽出する研究も複数存在する。評価文書は、直接的な評価を下す評価表現と、その評価対象である属性が共に現れる場合が多い。この評価対象を分析することは評判分析においても有益であると考えられる。

小林らは、評価文書から対象、属性、評価という 3 つ組を抽出することで、評価表現と属性を収集する手法を提案した[8]。この手法では、3 つ組を抽出するための共起パターンを定義し、既知の表現の周辺から新たな評価表現や属性の候補を収集する。しかし、この手法では抽出対象が 3 つ組の共起パターンを満たした文章だけであるため、評価文書によっては抽出が困難となる。一方、グエンらは評価表現を用いて、評価対象となる特徴語を複数のカテゴリに分類する手法を提案している[9]。語句同士の相互情報量を用いることで、特徴語を適切なカテゴリに分類ができると報告しているが、複数のカテゴリとそこに属する種語は予め人手で設定する必要がある、自動的に構築する上では問題がある。これに対し、予め種語集合を必要としない、評価表現辞書と特徴語の学習手法[10]や、評価要因に着目した評価表現の収集手法[11]も存在するが、いずれも収集した表現の肯定・否定を判定しないため、実際に評判分析に適用するには課題が存在する。

## 2.3 評判分析に関する研究

構築した辞書を用いた評判分析の手法としては、肯定語と否定語の出現割合に応じて評価を決定するものが存在する[12]。一方、分類器による機械学習を用いた手法も存在する。鈴木らは半教師付き学習の EM アルゴリズムとナイーブベイズ分類器を組み合わせる辞書の構築を行った。構築した学習モデルにより評価文書を肯定・否定・非評価の 3 クラスに分類することで評判分析を実現している[13]。

多面的な評判分析の研究には、コンテンツを見聞きした時に人々が感じる印象を、コンテンツそのものから抽出する研究が存在する[14]。しかし、この研究は[楽しい⇔悲しい]等の複数尺度における印象を分析するものであるため、商品の複数の評価軸を扱う本研究とは違った応用研究である。他にも、評価表現辞書の構築から分析までを一貫して行う研究としては、前川らが提案した、大規模ニュースを用いた株価収益率の予測が存在する[15]。この研究では、周辺文脈の情報を利用した手法[7]で収集した評価表現を利用して、ナイーブベイズモデルによる予測を行う。実験結果として、過去 1 日のニュース記事を用いて 5 日後の収益率を予測するモデルの精度が最も高くなることを報告しており、ニュースが発表されてから株価に反映されるまでのタイムラグを上手くとらえられると考察している。

## 2.4 既存研究における問題点

評価表現辞書を人手により構築する既存の手法[4][5]の問題点は、評価表現に対するラベル付けに膨大なコストがかかることにある。これに対して、評価表現辞書を自動的に構築するための複数の取り組み[6][7]では、それぞれ新語への対応が困難であることや、語彙情報を学習するためには大規模なコーパスが必要となることが課題としてあげられる[3]。加えて、これらの手法では評価文書を肯定・否定という1次元の尺度でしか評価できないため、ユーザに提示できる情報が不十分であると考えられる。本研究では、商品レビュー集合を学習コーパスとして用いることで、大規模コーパスの課題を解決できるため、周辺文脈の情報を利用した手法[7]をもとに手法の拡張を行う。

属性等の付加情報を抽出する手法[8][9][10]では、評価表現と共に評価対象となる属性の抽出を行うことができるが、評価表現の評価極性を分析していないため、実際に評判分析に用いるためには課題が存在する。また、収集した特徴語（属性）をカテゴライズするためには予め種語集合を用意する必要があり、複数の商品ジャンル（ドメイン）を対象とした場合に手間がかかってしまう。

また、評判分析では、文章に出現する肯定表現と否定表現の割合から評価を推測する手法[12]を紹介したが、単純な出現情報のみを用いた判定手法は精度面に問題があるほか、やはり肯定・否定という1次元の尺度でしか評価することが出来ない。一方で、分類器を用いた手法[13]では、それぞれの評価文書でどの表現が評価に寄与したかを判定することが困難であると考えられる。本研究では係り受け解析を用いて評判分析を行うことで、Turneyの手法[12]よりも高精度な推定が期待できるほか、分類器と異なり容易に評価に紐付く表現を提示することが可能である。

## 第3章 用語の定義

### 3.1 評価表現辞書

評価表現辞書は、肯定語と否定語を登録した語彙集である。文章中に出現する肯定語と否定語を手掛かりとして、その文章が肯定的なものか否定的なものかを推定するために用いる。評価表現辞書に関する語句の定義は以下の通りである。

評価表現： あるものを評価する語句である。評価表現の候補となる語句は、形容動詞語幹の名詞と形容詞である。

評価極性： 評価表現が示す肯定・否定の情報である。

評価極性値： 評価表現が示す評価の度合いを数値化したもの。 $[-1, 1]$ の連続値であり、評価極性値が正の場合は肯定、負の場合は否定の評価極性となる。

種表現： 学習の際に予め与えておく評価表現。商品ジャンルに依らず肯定・否定のいずれかを示すものを設定する。

今まで我々は、評価表現を肯定・否定という2値で表現してきた[16]。しかし、評価表現が示す肯定・否定の度合いは一樣ではないと考えられる。例として、「良い」という語句と「素晴らしい」という語句では、後者の方が肯定の度合いが強い場合が多いだろう。このような違いを表現するために、評価表現ごとに新たに評価極性値を算出する。評価極性値は連続値であり、この絶対値が大きいほど肯定・否定の度合いが強いということを表す。評価表現辞書の例を図3.1に示す。

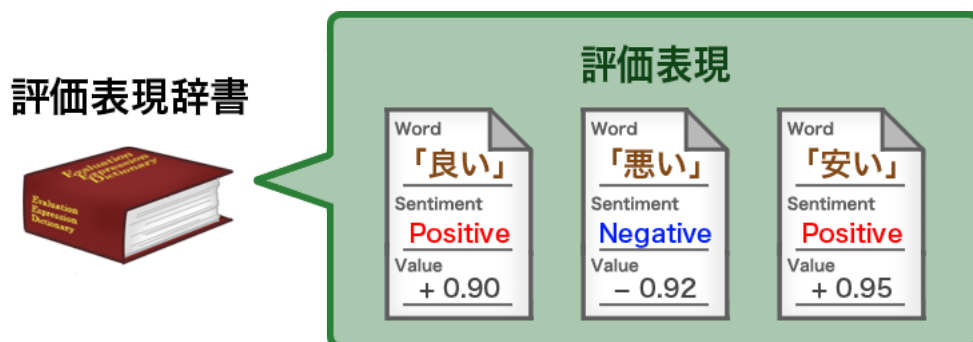


図 3.1 評価表現辞書

## 3.2 評価軸

評価軸は、「性能」や「価格」といった、対象の評価指標を表したものである。1.3 節で記述した通り、評価軸は特徴語の集合として表現される。

特徴語： ある評価軸に特有な、評価の対象となる語句である。特徴語の候補となる語句は一般名詞である。

代表評価表現： 評価軸内の特徴語と特に関係の強い評価表現である。

1つの評価軸は同じ意味を持つ特徴語集合によって形成され、この評価軸が複数構築される。それらは評価表現辞書と組み合わせることで、「商品の何が評価されているか」を推定するために利用する。評価軸の例を図 3.2 に示す。

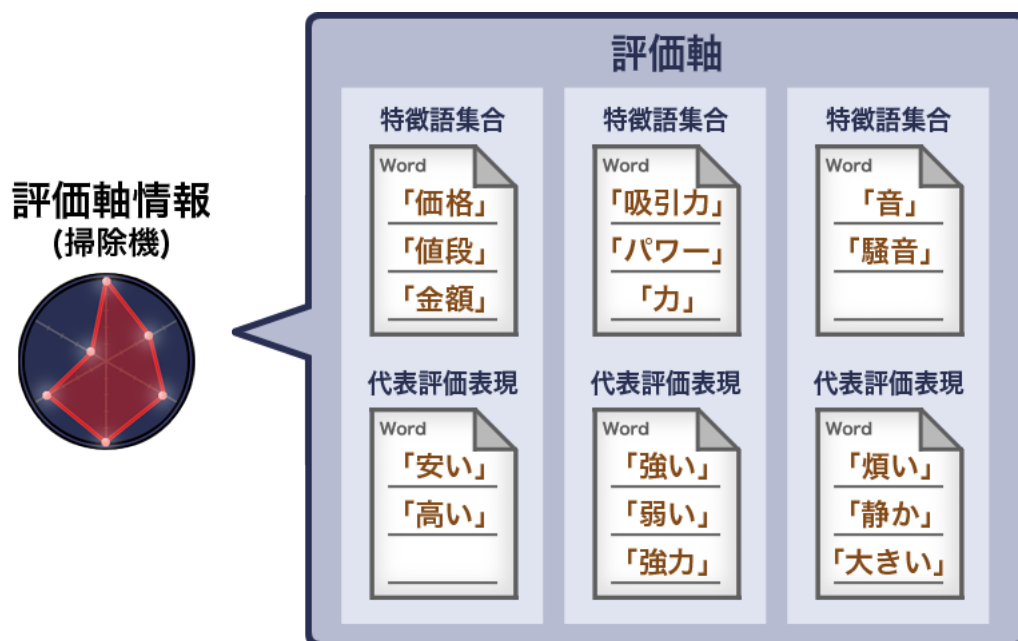


図 3.2 評価軸

## 3.3 商品ジャンル

評価軸と評価表現辞書は、商品ジャンル毎に構築する。商品ジャンルとは、商品を種類別にまとめる「掃除機」や「プリンタ」といった区分である。それぞれの商品ジャンルに特化した評価軸と評価表現辞書を構築することで、効果的な評判分析や情報推薦が可能となる。なお、評価実験における商品ジャンルは楽天市場<sup>1</sup>で定義されているものを利用する。

<sup>1</sup> <http://www.rakuten.co.jp/>

## 第4章 語彙情報の学習

本研究の提案手法は、実際のレビューからその商品ジャンルに適した評価軸と評価表現辞書を構築する「語彙情報の学習」と、それをを用いて商品レビューを要約する「評判分析」に大別される。本章では、レビューからの「語彙情報の学習」について提案する。

### 4.1 学習手順

ある商品ジャンルのレビュー集合と少数の種表現を入力として、そのジャンルに適した評価軸と評価表現辞書を構築する。図4.1はその学習フローである。

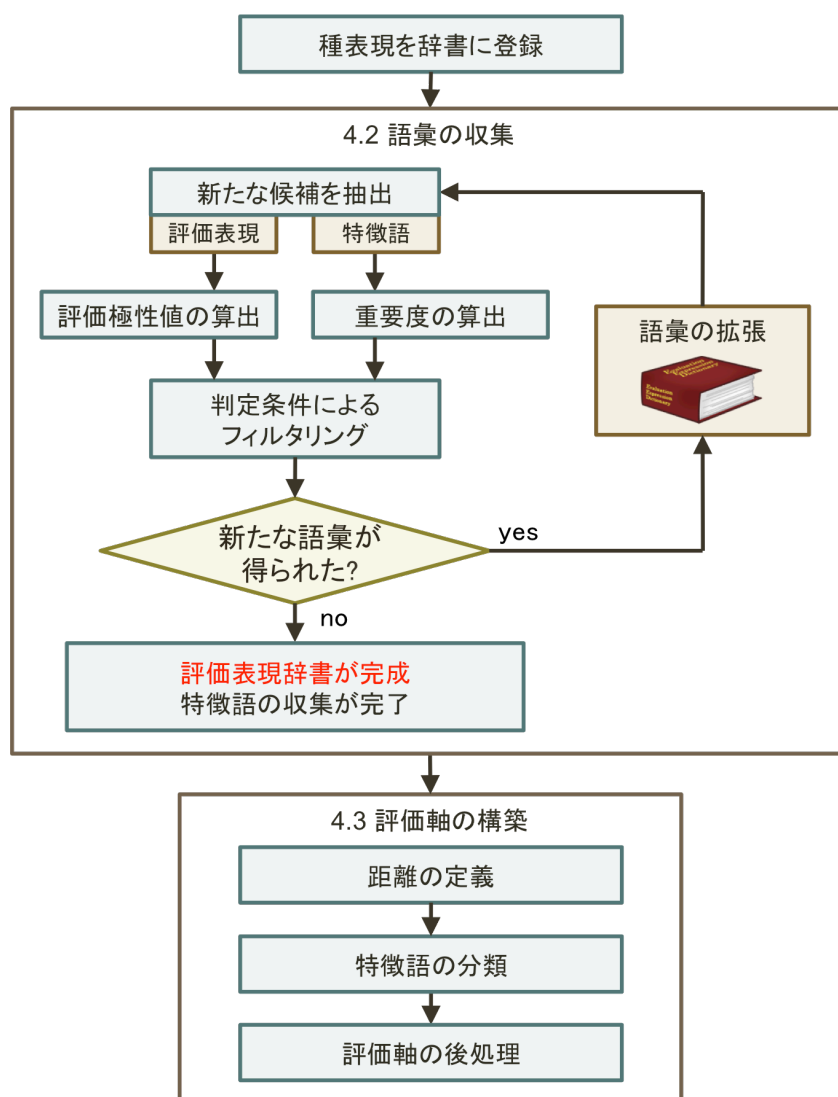


図 4.1 語彙情報の学習フロー

種表現は、ジャンルに依らず肯定・否定が一意に定まる語句を人手で選出する。その際の評価極性値は、肯定語は+1, 否定語は-1とする。種表現の設定にそれ以外の条件はなく、数は高々10語程度である(6章の実験では2語とした)。初期値として種表現を辞書に登録し、「4.2 語彙の収集」のステップでコーパスの商品レビューから新たな語彙を獲得していく。この工程で評価表現の収集が完了、すなわち評価表現辞書が完成する。同じく特徴語の収集も完了するが、さらに分類により、類似した意味を持つ集合にまとめる必要がある。この工程が「4.3 評価軸の構築」である。

## 4.2 語彙の収集

辞書の語彙をもとに、新たな評価表現と特徴語を収集していく。語彙を収集する手法は、那須川らの提案した周辺文脈の情報を利用した手法[7]をもとにしているが、以下の2つの点で拡張を行っている。

1. 評価表現の評価極性を連続値(評価極性値)で表現する。
2. 評価表現だけでなく、評価対象となる特徴語も抽出する。

候補の獲得では、レビュー中に出現する既知の評価表現の周辺から、新たな候補を抽出していく。その後、判定条件によるフィルタリングを行うことで、辞書に登録する語彙を決定する。以上の手順を、新たな語彙が得られる限り繰り返すことで、ブートストラップ的に辞書を拡張することができる。また、繰り返しの度に、既知の評価表現の評価極性値を更新することで、それぞれの評価表現に適した肯定・否定の度合いを定めることが出来る。従って、種表現であっても最終的な評価極性値は+1, -1以外の値を取りうる。

新たな語彙が得られなくなった時点で「語彙の収集」が完了し、「評価軸の構築」の手順に移る。

### 4.2.1 評価表現候補の抽出

文章中に既知の評価表現が現れた場合、次の条件を満たす並列な文節から評価表現候補を抽出する。

- ・ 接続詞を伴って隣接する前の文章の主節
- ・ 同一文中の係り受け関係にある文節
- ・ 接続詞を伴って隣接する後の文章の主節

これらの具体例は4.2.3節で記述する。なお、既に登録された評価表現であっても評価極性値の更新のために抽出を行うが、同じ表現同士では抽出しない。



また、評価表現候補を抽出する際には、抽出元の評価表現を手掛かりに、評価極性値を推定する。候補の抽出毎に抽出スコアを算出し、全レビューから候補を抽出した後、スコアを集計して評価極性値を決定する。既知の評価表現  $e$  によって評価表現候補  $x$  を抽出した際の抽出スコア  $extract(e, x)$  は以下の式で算出する。

$$extract(e, x) = value_e \cdot conjunction(e, x) \cdot reverse(e) \cdot reverse(x) \quad (4.1)$$

但し、

$$conjunction(x, x') = \begin{cases} +1 & (\text{表現}x\text{と表現}x'\text{が順接である}) \\ -1 & (\text{表現}x\text{と表現}x'\text{が逆接である}) \end{cases}$$

$$reverse(x) = \begin{cases} +1 & (\text{表現}x\text{の評価極性が文中で反転していない}) \\ -1 & (\text{表現}x\text{の評価極性が文中で反転している}) \end{cases}$$

であり、 $value_e$  は評価表現  $e$  の評価極性値である。なお、特に既知の評価表現を指定しない場合は  $extract(x)$  と表記する。

これは、既知の評価表現と候補が順接関係にあれば同じ極性、逆接関係にあれば逆の極性だと推測する事を意味している。「評価極性が反転する」とは、極性反転子「ない」等が評価表現に続くことで、本来とは逆の極性を示すことを指す。反転情報は文節内の極性反転子の数に応じて決定する。「わからなくもない」といった表現に対応するため、文節内の極性反転子が奇数個であれば「反転している」、偶数個であれば「反転していない」と見なす。極性反転子や逆接の接続詞は予め人手で設定しておく。

#### 4.2.2 特徴語候補の抽出

特徴語候補の抽出も係り受け関係をもとに行う。既知の評価表現と係り受け関係にある文節を探索し、候補となる語句があれば抽出する。特徴語を抽出する際には、「評価軸の構築」で特徴語を分類するため、抽出に寄与した評価表現も同時にカウントしていく。

しかし、全ての係り受け関係を共起関係として扱うのは望ましくない。例として、「音はうるさいけど、十分な吸引力だ」という文章を考える。特徴語の分類に用いる共起ペアとしては、(音, うるさい) と (吸引力, 十分) の2つが抽出される事が望ましいが、単純な係り受け関係を用いると (吸引力, うるさい) という組も抽出してしまう。本研究ではこれに対応するため、特徴語と評価表現の共起関係として扱うのは、「係り受け関係にある文節のうち、接続詞または接続助詞を伴わずに係るもの」とした。

### 4.2.3 候補の抽出例

図 4.2 に肯定の種表現として「良い」を設定した場合の抽出例を示す。評価表現候補を下線、特徴語候補を二重線、接続詞または接続助詞を破線で表している。

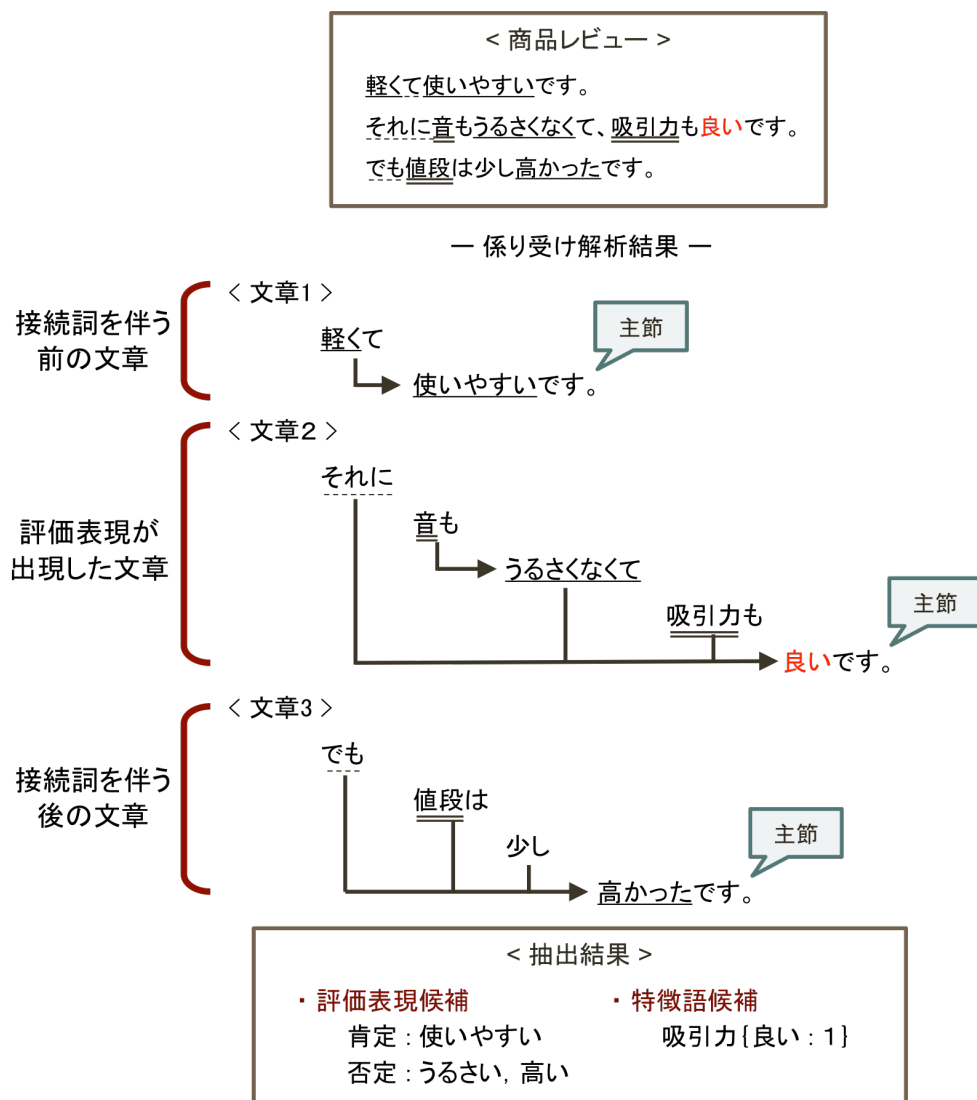


図 4.2 候補の抽出例

評価表現の現れた文章においては、係り受け関係から「吸引力」が「良い」と共起する特徴語候補として抽出される。次に、同一文中からは「うるさい」が、極性反転子「ない」を考慮して、否定表現として抽出される。また、文章の先頭に接続詞「それに」があるため、この文章は一つ前の文章と順接関係にあると推測され、「使いやすい」が肯定表現として抽出される。一方で、一つ後の文章は先頭に接続詞「でも」がある。従って、評価表現の現れた文章とは逆接関係にあると推測され、「高い」が否定表現として抽出される。

#### 4.2.4 評価表現候補の評価極性値の算出

評価表現の評価極性値は、4.1式で得た抽出スコアから算出する。まず、それぞれの評価表現候補  $x$  について、以下のスコア  $score(x)$  を計算する。

$$score(x) = \frac{\sum extract(x)}{\sum |extract(x)|} \quad (4.2)$$

評価表現候補  $x$  が新たな語彙である場合は、このスコアを評価極性値とする。既知の評価表現を更新する際は、パラメータ  $\alpha$  ( $0 < \alpha < 1$ ) を用いて、以下の式で求める。

$$value'_x = score(x) \cdot \alpha + value_x \cdot (1 - \alpha) \quad (4.3)$$

この  $\alpha$  の値は予め設定しておく。

#### 4.2.5 特徴語候補の重要度の算出

特徴語候補が特徴語かどうかを判定するために、特徴語の重要度を定義する。重要度を決定する要素としては以下の3つを設定した。

1. 学習コーパスから特徴語候補として抽出された回数
2. 抽出時に特定の助詞（係助詞，格助詞）を伴った確率
3. コーパス全体で、評価表現候補と係り受け関係にある確率

助詞との共起確率を要素としているのは、特徴語は主語や目的語として現れる可能性が高く、候補として抽出された特徴語の前後の品詞を考慮することが有益だという報告に基づくものである[17]。3つ目の項目は、形容詞（形容動詞語幹）との共起のしやすさを表しており、人物を始めとする汎用的な語句のフィルタリングに寄与すると期待するものである。特徴語候補  $y$  の重要度  $weight_f$  は以下の式で与える。

$$weight_y = \sqrt{extract\_num(y) \cdot rate\_particle(y) \cdot rate\_EE(y)} \quad (4.4)$$

但し、

- $extract\_num(y)$  : 特徴語候補  $y$  の抽出回数  
 $rate\_particle(y)$  : 特徴語候補  $y$  が抽出時に特定の助詞を伴った確率  
 $rate\_EE(y)$  : 特徴語候補  $y$  が評価表現候補と係り受け関係にある確率

である。

#### 4.2.6 判定条件によるフィルタリング

抽出した評価表現候補と特徴語候補にはノイズが多く存在すると考えられるため、それぞれ判定条件を用いてフィルタリングを行う必要がある。

評価表現の判定では、候補として抽出された回数と、4.2.4節で算出した評価極性値による閾値判定を行う。両方を満たした語句を、評価表現として評価極性値と共に評価表現辞書に登録する。図4.3は評価表現判定のフローである。

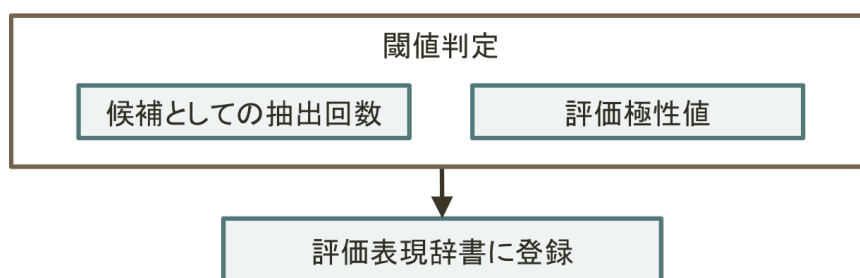
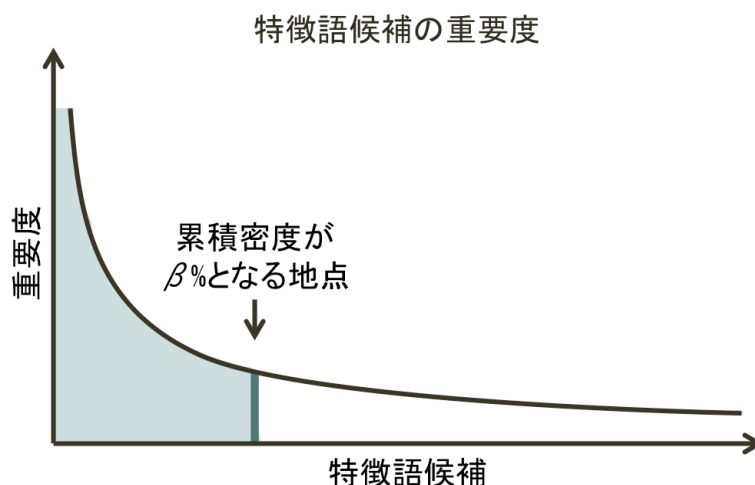


図 4.3 評価表現判定のフローチャート

特徴語候補に対して行う閾値判定は、「格助詞を伴う抽出確率」と「係助詞を伴う抽出確率」である。4.4式で算出した重要度は、格助詞と係助詞のいずれかと共起する確率が高ければ、高い値が得られる可能性がある。しかし、特徴語候補として抽出される一般名詞の中には「割」や「全体的」等のノイズとなる語句が存在し、それらは「割と～」や「全体的に～」といったような形で、格助詞と係助詞の一方と極端に共起する傾向を持つ。そのため、片方の助詞との共起が極めて低い候補は予め除去しておく。なお、この閾値は誤って特徴語となるべき語句を除去しないよう、十分に低い値を設定するものとする。

評価軸として扱うべき特徴語は、多くのレビューで評価されており、ユーザの関心が強いと考えられる語句である。そのため、扱う商品ジャンルによって異なる、抽出される特徴語候補の種類や数、評価表現数に左右されずに、重要な語句だけを選別する仕組みが必要である。単純な閾値判定だけでは、抽出される評価表現の影響を受けやすいと考えられるため、これに加えて、抽出した候補全体の統計値を用いたフィルタリングを行う。式4.4の重要度は、抽出した全ての一般名詞に対して計算されるため、その大部分はノイズである。一方で、商品レビュー内で頻繁に評価される語句は高い値となるため、特徴語候補全体での重要度の分布はロングテールとなることが想定される。そこで、重要な特徴語とそれ以外を切り分けるために、図4.4で示すように、重要度において全体の $\beta\%$ を占める上位語句だけを特徴語として採用する。



上位  $\beta\%$  を占める候補を特徴語とする(青色部)

図 4.4 重要度による特徴語判定の例

ここで採用された特徴語が、「評価軸の構築」でクラスタリングの対象となる。式 4.4 より、抽出回数が多いものが相対的に高い重要度となりやすく、商品ジャンルに対応する評価軸を構築する際には上位の表現をクラスタリング対象とした方が良い結果を得やすい。

しかし、ロングテール部分にも特徴語とすべき表現は少なからず存在すると考えられる。特定の特徴語と類似した意味を持つが、出現頻度が極端に低いものは重要度が低くなってしまいうため、ノイズとみなされてしまう。これは評価軸を構築する上ではさほど問題ではないが、ひとつひとつの商品を対象とする評判分析では、なるべく多くの特徴語が存在する方が有利である。そのため、ロングテール部分の表現は「準特徴語」とし、特徴語のクラスタリング後に評価軸に分配する。これについては 4.3 節で説明する。

特徴語判定のフローを図 4.5 に示す。

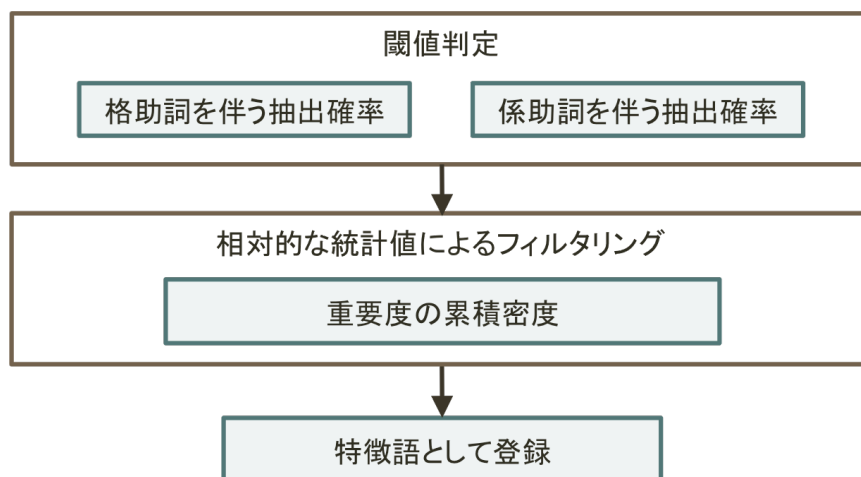


図 4.5 特徴語判定のフローチャート

### 4.3 評価軸の構築

学習によって得られた特徴語は、類似した意味を持つ特徴語同士が集合することで評価軸となる。正しく評価軸を構築するには分類を正しく行う必要があるため、類似した意味の特徴語同士が近くなるような距離を定義することが重要である。

#### 4.3.1 特徴語間の距離の定義

本研究では特徴語の意味的な類似度を表現するために TF-IDF を流用する。TF-IDF とは本来「文書集合において、ある単語が各文書内でどれだけ重要であるか」を表す尺度であり、以下の式で表される。

$$tfidf_{i,j} = tf_{i,j} \cdot idf_i$$

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (4.5)$$

$$idf_i = \log \frac{|D|}{|\{d:d \ni t_i\}|}$$

$n_{i,j}$  は単語  $t_i$  の文書  $d_j$  における出現回数で、 $|D|$  は文書の総数、その分母は単語  $t_i$  が登場する文書数を表している。 $tf$  は Term Frequency の略であり、単語の出現頻度を表している。一方、 $idf$  は Inverse Document Frequency であり、逆文書頻度と呼ばれる。「今日」等のどんな文書にでも頻繁に登場するような汎用語は、 $tf$  の値は大きくなるが、 $idf$  の値は小さくなるため、TF-IDF の値は大きくなる。従って、特定の文書内で頻繁に現れる、重要な単語に対してのみ大きな値が付与される仕組みになっている。

この TF-IDF は文書分類を目的として用いることが多いが、この TF-IDF を「特徴語と評価表現の関連度」を表す尺度として用いる。これは「同じ評価軸に属する特徴語は、類似した評価表現と共起する可能性が高い」という仮定に基づいている。分類対象は特徴語であるため、本来の TF-IDF との各要素の対応表は表 4.1 のようになる。

表 4.1 本来の TF-IDF との対応表

本来の要素	対応する要素
単語 $t_i$	評価表現 $e_i$
文書 $d_j$	特徴語 $f_j$

よって、本研究における TF-IDF は次式で表される。

$$tfidf_{i,j} = tf_{i,j} \cdot idf_i$$

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (4.6)$$

$$idf_i = \log \frac{F}{F_i}$$

$n_{i,j}$ は評価表現 $e_i$ と特徴語 $f_j$ が係り受け関係にあった回数であり、 $F$ は特徴語の総数、 $F_i$ は評価表現 $e_i$ と共起する特徴語の数である。よって TF-IDF は、「ある特徴語とある評価表現の相関関係」の指標となる。

以下に TF-IDF の計算例を示す。「4.2 語彙の収集」によって、以下の語彙と係り受け回数を得たとする。

表 4.2 獲得した語彙

特徴語	値段, 価格, 音
評価表現	良い, 高い, うるさい

$$\begin{array}{c} \text{値段} \quad \text{価格} \quad \text{音} \\ \text{良い} \quad \left( \begin{array}{ccc} 5 & 3 & 0 \\ 10 & 8 & 0 \\ 0 & 0 & 10 \end{array} \right) \\ \text{高い} \\ \text{うるさい} \end{array}$$

これに対し、上記の TF-IDF の式を適用することで、以下のような TF-IDF 行列が完成する。

$$\begin{array}{c} \text{値段} \quad \text{価格} \quad \text{音} \\ \text{良い} \quad \left( \begin{array}{ccc} 0.14 & 0.11 & 0 \\ 0.27 & 0.29 & 0 \\ 0 & 0 & 1.10 \end{array} \right) \\ \text{高い} \\ \text{うるさい} \end{array}$$

これにより各特徴語と各評価表現の関連度を定義することができる。ここで特徴語を、評価表現数を次元とする空間の点とみなすことで、特徴語間の距離を定義できるため、クラスタリングによる特徴語の分類が可能となる。なお、実際に用いる値は、TF-IDF 値を特徴語ごとに正規化したものである。

### 4.3.2 特徴語の分類

特徴語の分類は、クラスタリングによって実現する。特徴語を分類したクラスタひとつひとつが評価軸となる。特徴語のクラスタリングには、教師なし学習である **k-medoids** 法を用いる。**k-medoids** 法は、同じく教師なし学習である **k-means** 法に類似したクラスタリング手法である。**k-means** 法では、クラスタの重心であるセントロイドをクラスタの代表点とし、最も近いクラスタへのデータの再分類とセントロイドの更新を繰り返す。一方、**k-medoids** 法では、クラスタの代表点に **medoid** というデータを用いる。この **medoid** は各クラスタに属するデータから 1 つずつ選出され、自分以外のデータとの距離の総和が最小となるものが選ばれる。それぞれの手法の代表点決定の例を図 4.6 に示す。クラスタ内の平均点を算出する **k-means** 法と比較して外れ値の影響が小さいという利点があり、一部の語句が高い重要度を持つ本研究の分類対象に対しては **k-medoids** 法の方が有効であると考えられる。しかし、これらの手法にはクラスタ数を予め指定する必要があるという問題点がある。商品ジャンルによって評価軸の数は異なると考えられるため、クラスタ数を自動決定する手法が必要である。

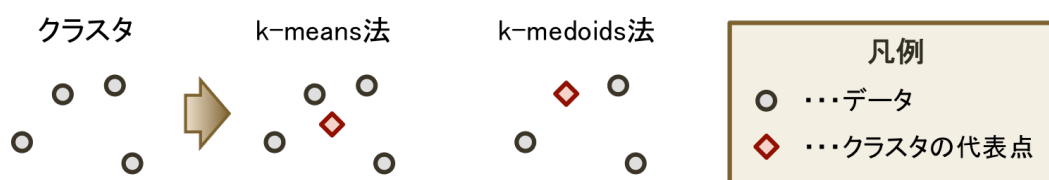


図 4.6 クラスタの代表点決定法

クラスタ数を自動決定するための手法は複数提案されている。**X-means** 法は、再帰的な **k-means** 法によって最適な分割数とクラスタを決定する手法である[18][19]。十分に小さい分割数から **k-means** クラスタリングを実行し、分割結果のクラスタに対して、分割が妥当だと判断される限り、**2-means** 法による分割を繰り返すというものである。一方、**Gap** 統計量を用いた手法も存在する[20]。**Gap** 統計量は、一様分布から生成されたデータの分類結果と実データの分類結果のクラスタ内距離平均の差として与えられ、これが大きいほど、実データのクラスタリング結果が密であると言える。クラスタ数を  $k = 1, 2, \dots$  と増やしながらか **Gap** 統計量を計算し、局所最適な最小の  $k$  の値を選択するというものである。後者の手法は反復回数が多いため処理時間は長くなるが、**2-means** 法を繰り返す **X-means** 法よりも分類結果が安定すると考えられる。本研究では、学習により一度評価軸を構築すれば、評判分析に活用することができるため、特徴語の分類にかかる時間はさほど問題にならない。したがって、**Gap** 統計量を用いてクラスタ数を自動決定する手法を採用する。



### 4.3.3 評価軸の後処理

Gap 統計量で決定した  $k$  の値によるクラスタリングは、クラスタ内の凝集性が高くなる一方で、類似した特徴語がない場合は要素数が極端に小さいクラスタが形成される可能性がある。そのような評価軸は、ノイズとみなせるならば除去することが好ましい。そこで、クラスタリング後の評価軸に対して、以下の手順で判定を行う。

1. 特徴語全体の重要度の総和を計算する。
2. 評価軸ごとの重要度の合計を計算する。
3. 特徴語全体の総和に占める割合が閾値を下回る評価軸はノイズとして除去する。

商品ジャンルにおいて重要な評価軸はレビュー中で頻繁に評価されるため、それを構成する特徴語の重要度は大きくなる傾向がある。従って、この重要度を用いてノイズとなる評価軸を除去することで、重要な評価軸のみを残すことが出来ると考えられる。

次に、4.2.6 節で「準特徴語」とした表現の、評価軸への分配を行う。これは、商品の評判分析ではなるべく多くの特徴語が存在する方が有利なためである。ある語句の出現頻度がコーパス全体では少なくとも、商品ごとのレビューでは頻繁に出現する可能性があり、そのような語句をカバーすることは重要である。

具体的には、それぞれの準特徴語について、以下の処理を行う。

1. 最も距離が近い評価軸との距離を  $d_1$  とする。
2. 2 番目に距離が近い評価軸との距離を  $d_2$  とする。
3.  $d_1/d_2$  が閾値を下回れば、準特徴語を最も近い評価軸に分配する。

特定の評価軸に属さない語句は、複数の評価軸と同程度の距離となる場合があるため、いずれかの評価軸と極端に近い場合に、準特徴語の分配を行う。

最後に、評価軸ごとに代表評価表現を決定する。第 5 章で説明する評判分析の手法では、主に特徴語と評価表現の係り受け関係に基づき、評価軸ごとの評価を決定する。しかし、レビューの中には評価表現単体で評価軸を表しているケースも多数存在する。そこで、そのような評価箇所をカバーするために評価軸ごとの代表評価表現を決定し、特徴語を伴わない箇所の評価抽出の改善を図る。代表評価表現は、評価軸の代表点の評価表現ベクトルから決定する。合計が 1 となるように正規化したベクトル上で、値が閾値以上のものを代表評価表現だと判定する。複数の評価軸において評価表現が競合する場合には、より大きい値である評価軸の代表評価表現として採用する。なお、具体的な評価抽出方法については第 5 章で紹介する。

## 第5章 レーダーチャート生成システム

### 5.1 システム構成

本研究のシステム構成を図 5.1 に示す。図の上部は 4 章で述べた語彙情報の学習システムであり、その結果として、ジャンル毎の評価軸情報と評価表現辞書が得られる。下部が評判分析システムであり、ユーザが商品を指定すると、レビューを要約してユーザに提示する。この際に、商品が属するジャンルの評価軸情報を利用した分析を行うことで、多面的な評価が可能になり、レーダーチャートによる分析結果を提示することができる。

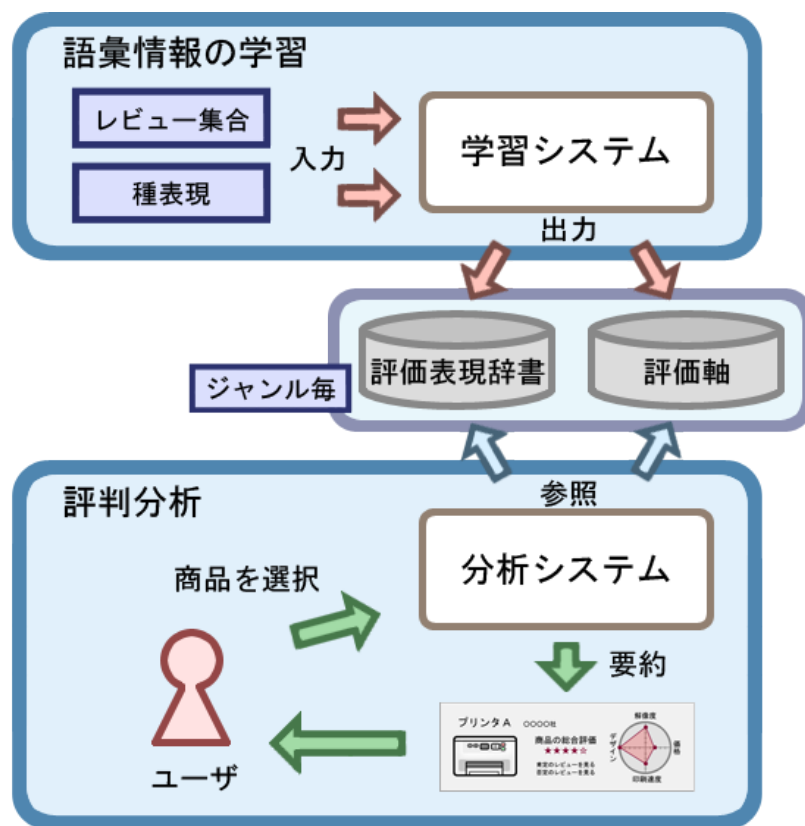


図 5.1 評判分析システムへの応用

評判分析も語彙情報の学習と同じく係り受け解析を利用し、評価表現の極性や反転情報をもとに、評価値の計算を行う。本研究の評判分析では、レビュアーが商品を肯定的に評価しているかという商品の「総合評価」と、商品ジャンルに応じた「評価軸ごとの評価」の二つを知ることができる。

## 5.2 レーダーチャート生成の手順

商品を選択してから，レーダーチャートを生成するまでの手順を図 5.2 に示す．

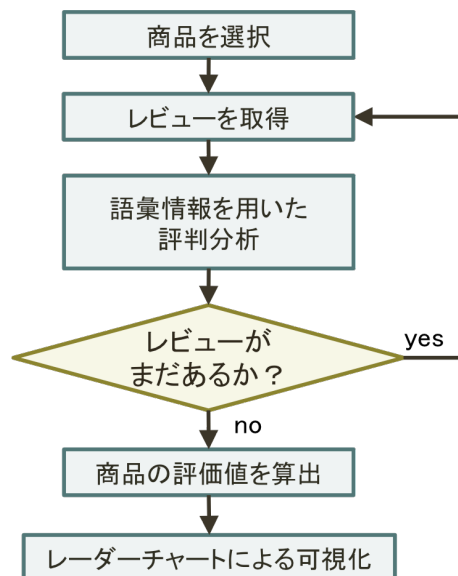


図 5.2 レーダーチャート生成までのフローチャート

当該商品の全ての商品レビューに対して，学習した語彙情報を用いた評判分析を行い，各レビューの評価値を算出する．この評価値を標本値として，ジャンル内の評価値の平均と標準偏差から，偏差値を算出する．この偏差値を集約し，商品の評価値として出力する．

これらの評価値は総合評価や評価軸ごとに個別に求める．偏差値を用いることで，ある商品のある評価軸が，ジャンル内で相対的にどの程度の評価をされているか知ることが可能となる．

## 5.3 各レビューの評価値の算出方法

総合評価の分析には，文章中の主節の評価表現を利用する．これは「文章全体の評価極性は主節の極性と一致する可能性が高い」ためである．文章毎の主節に評価表現があれば，反転情報を加味してその評価極性値を加算していく．各レビューの評価値は文章毎の評価極性値の和となる．

一方，評価軸ごとの評価は，基本的に構築した評価軸に属する特徴語をもとに，係り受け関係を手掛かりとして行う．文中で「接続詞や接続助詞を伴わずに係り受け関係にある」評価表現と特徴語が存在する場合，その特徴語が属する評価軸の評価として，反転情報を加味した評価極性値をカウントしていく．各レビューの評価値は，評価軸ごとの評価極性値の和である．

また、評価軸ごとの評価は、代表評価表現を用いることでも抽出する。評価表現の中には、特徴語を伴わずに特定の評価軸を表現するものがしばしば見られる。例として「安い」という評価表現は、「価格」といった特徴語を伴っていないとしても、ほぼ確実に金額についての評価を示す語句である。このような代表評価表現がレビュー中でどの特徴語とも係り受けせずに出現した場合は、例外的にその表現が属する評価軸の評価としてカウントを行うものとする。

なお、上述の各レビューの評価値は、偏差値を計算する際の標本値となるが、平均や分散を求めるための標本値として考慮されるのは評価がなされていたレビューのみである。つまり、ある評価軸については、その軸が評価されたレビュー集合を対象として、偏差値の計算を行う。

図 5.3 は評判分析の処理フローである。

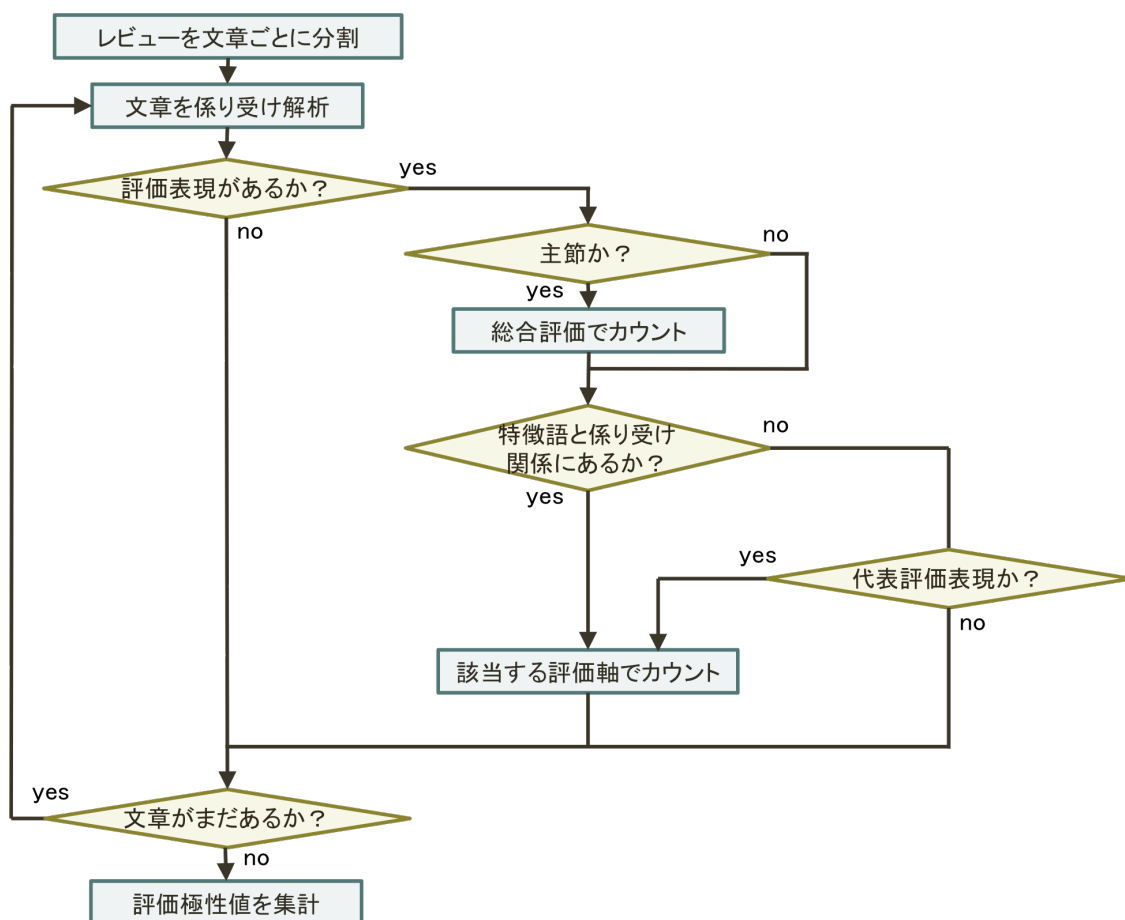


図 5.3 評判分析のフローチャート

評判分析の例を以下に示す．語彙情報の学習により，表 5.1 の評価表現辞書と表 5.2 の評価軸を構築したとする．

表 5.1 獲得した評価表現辞書の例

評価表現	評価極性	評価極性値
満足	肯定	+0.95
強い	肯定	+0.90
うるさい	否定	-0.92

表 5.2 獲得した評価軸の例

	特徴語群	代表評価表現
評価軸 1	音，動作音	うるさい
評価軸 2	吸引力，力	強い
評価軸 3	価格，値段，金額	

ここで，「少しうるさいけど，吸引力は強いし満足です」という商品レビューが与えられたとする．その係り受け解析結果から評価箇所を抽出した例を図 5.4 に示す．

— 係り受け解析結果 —

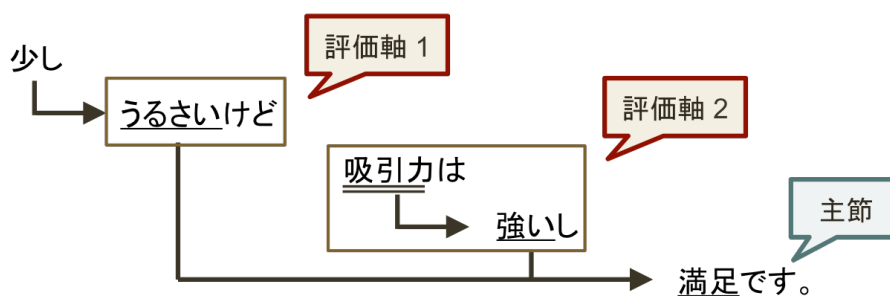


図 5.4 評価箇所の抽出例

この文章の主節は「満足です」であり，肯定表現「満足」を手掛かりに，この文章は肯定評価されていると推定される．その際に，「満足」の評価極性値 (+0.95) をこの商品レビューにおける総合評価として加算する．

一方で，肯定表現「強い」は特徴語「吸引力」と係り受け関係にある．「吸引力」は評価軸 2 に属しているため，この商品レビューにおける評価軸 2 の評価として「強い」の評価

極性値 (+0.90) を加算する。また、否定表現「うるさい」はどの特徴語とも係り受け関係にないが、評価軸 1 の代表評価表現である。そのため、この商品レビューにおける評価軸 1 の評価として「うるさい」の評価極性値 (-0.92) を加算する。

このようにして、レビュー中の全ての評価表現について、主節の表現であるか、特徴語と係り受け関係にあるか、代表評価表現であるかを判定していく。この例では、評価軸 3 については評価されておらず、このレビューに対する評判分析結果は表 5.2 のようになる。

表 5.2 レビューにおける評価値の計算結果の例

	評価値	評価レビュー数
総合評価	+0.95	+1
評価軸 1	-0.92	+1
評価軸 2	+0.90	+1
評価軸 3	±0	+0

この計算を商品の全レビューに対し行うことで、商品の最終的な評価値が決定される。なお、表中の評価レビュー数は、平均や標準偏差を求める際の母数としてカウントするかどうかを指しており、総合評価や特定の評価軸が多重に評価されている場合も+1 である。

## 第 6 章 評価実験

提案手法が幅広いジャンルに対して有効であることを示すために、評価実験を行った。実験においては、複数の商品ジャンルのレビュー集合に対して、「語彙情報の学習」と「評判分析」を行い、その結果について考察する。

### 6.1 実験条件

#### 6.1.1 楽天データセット

語彙情報の学習では、学習用コーパスとしてジャンル毎の商品レビュー集合を必要とする。実験では、楽天株式会社と国立情報学研究所が公開している「みんなのレビュー・ロ

表 6.1 「みんなのレビュー・ロコミ情報」のデータ構造

分類	カラム	内容についての補足
投稿者情報	投稿者	「user1」のようにマスクしたユーザ名
	年齢	「10 歳代」等
	性別	[0:男 1:女 2:不明]
商品情報	商品コード	商品の主キー
	商品名	—
	店舗名	—
	商品 URL	—
	商品ジャンル ID	商品ジャンルの主キー
	商品価格	—
レビュー情報	購入フラグ	購入者かどうか
	内容	「実用品・普段使い」等の文字列
	目的	「自分用」等の文字列
	頻度	「はじめて」等の文字列
	評価ポイント	0 から 5 までの 6 段階評価
	レビュータイトル	—
	レビュー内容	—
	レビュー登録日時	—

コミ情報」というデータセットを使用する。これには楽天市場の商品のユーザレビュー情報が収められており、2010年から2012年までの3年間分のレビューが公開されている。実際に格納されているデータ構造を表6.1に示す。本研究ではこのデータセットから、「レビュータイトル」と「レビュー内容」のカラムを合わせて、一つのレビューとして扱う。

評価実験で扱うジャンルは表6.2の通りである。商品ジャンルは楽天市場において木構造で定義されており、これをもとに商品进行分类してそれぞれのコーパスを生成している。但し、「掃除機」ジャンル内の「ロボット掃除機」のように、ジャンルの構造上ノイズとなり得ると判断したものについては除外している。また、大量の商品レビューの中には、同一ユーザによる同一内容のものもしばしば見られるため、同一内容のレビューは1つしか登録しないようにした。

表 6.2 商品ジャンル毎のコーパス情報

ジャンル名	レビュー数
掃除機	78504
スーツ	93475
和菓子	78847

### 6.1.2 閾値

第4章で述べた、語彙情報の学習に関するパラメータはそれぞれ表6.3で示すように設定した。これらは、多くの商品ジャンルに対して安定した結果を出せるよう、経験的に選択した値である。

表 6.3 評価実験におけるパラメータ

	パラメータ	実験における値
評価表現	抽出回数の閾値	10
	評価極性値の絶対値の閾値	0.85
	評価極性値のパラメータ $\alpha$	0.8
特徴語	特定の助詞との共起確率の閾値	10%
	重要度による判定閾値 $\beta$	50%
評価軸	ノイズの評価軸とみなす閾値	0.05
	準特徴語を分配する $d_1/d_2$ の閾値	0.5
	代表評価表現の閾値	0.1



## 6.2 評価表現辞書の構築

### 6.2.1 語彙の収集結果

肯定表現「満足」と否定表現「不満」の2語を種表現として、評価表現辞書の構築実験を行った。それぞれの商品ジャンルにおける獲得語彙数を表6.4に示す。

表 6.4 商品ジャンル毎の獲得語彙数

ジャンル名	肯定語	否定語
掃除機	67	24
スーツ	48	11
和菓子	51	14

収集した評価表現を表6.5から表6.7に示す。

表 6.5 「掃除機」の評価表現辞書の構築結果

	評価表現
肯定表現	ありがたい, いい, おしゃれ, お洒落, かつこいい, かつこよい, かわいい, きれい, すばらしい, ばっちり, ぴったり, よい, イイ, オシャレ, カッコイイ, キレイ, コンパクト, シンプル, スムーズ, スリム, ハンディ, パワフル, ラッキー, 丁寧, 不要, 便利, 優秀, 十分, 可愛い, 可能, 大丈夫, 好き, 嬉しい, 安い, 安価, 安心, 小さい, 小型, 強い, 快適, 手ごろ, 手軽, 手頃, 抜群, 早い, 易い, 有り難い, 楽, 楽しい, 楽ちん, 気持ちいい, 満足, 簡単, 素敵, 素晴らしい, 綺麗, 良い, 薄い, 見やすい, 赤い, 軽い, 迅速, 透明, 速い, 静か, 面白い, 高性能
否定表現	うるさい, ごつい, しんどい, つらい, でかい, 不便, 不明, 不満, 億劫, 固い, 大きい, 大変, 太い, 残念, 無理, 煩い, 痛い, 辛い, 遅い, 邪魔, 重い, 重たい, 難しい, 高い

表 6.6 「スーツ」の評価表現辞書の構築結果

	評価表現
肯定表現	ありがたい, いい, うれしい, おしゃれ, お洒落, かっこいい, かっこよい, かわいい, きれい, すばらしい, はやい, ばっちり, やすい, やわらかい, よい, イイ, オシャレ, カッコイイ, キレイ, 丁寧, 丈夫, 便利, 充分, 可愛い, 大丈夫, 嬉しい, 安い, 安心, 快適, 手ごろ, 手頃, 早い, 暖かい, 有難い, 柔らかい, 楽, 気持ちいい, 渋い, 満足, 素敵, 素晴らしい, 良い, 親切, 豊富, 軽い, 迅速, 速い, 駄目
否定表現	うすい, 不安, 不満, 地味, 残念, 派手, 短い, 薄い, 遅い, 重い, 雑

表 6.7 「和菓子」の評価表現辞書の構築結果

肯定表現	うまい, うれしい, おいしい, かりかり, かわいい, きれい, すごい, たまらない, ぴったり, みずみずしい, めずらしい, もったいない, やわらかい, よい, カリカリ, ヘルシー, 丁寧, 丁度, 便利, 優しい, 冷たい, 凄い, 可愛い, 大きい, 大好き, 安い, 安心, 幸せ, 懐かしい, 手ごろ, 手頃, 早い, 有名, 有難い, 柔らかい, 楽しい, 満足, 濃い, 珍しい, 甘酸っぱい, 程よい, 綺麗, 美味しい, 美味しい, 良い, 親切, 軟らかい, 軽い, 迅速, 面白い, 香ばしい
否定表現	くどい, しつこい, 不安, 不満, 割高, 大変, 小さい, 小さめ, 微妙, 普通, 残念, 物足りない, 遅い, 高い

以上の構築結果からは、各商品ジャンルに特有な評価表現が複数見られる。「掃除機」のジャンルでは、「強い」や「パワフル」といった、吸引力に関するもの、「静か」や「煩い」等の音に関する評価表現が含まれる。「スーツ」のジャンルでは、「柔らかい」や「薄い」等の生地に関するものの他に、「渋い」や「地味」、「派手」といった評価表現も見て取れる。「和菓子」のジャンルでは「美味しい」や「甘酸っぱい」、「香ばしい」のように、味や風味に関する語句が抽出できている。

## 6.2.2 評価極性値の評価

収集した評価表現には、4.2.4 節で記述したように評価極性値を付与している。次に、この評価極性値の妥当性について考察する。比較対象として、「評価表現が商品レビュー内で実際に肯定語・否定語として使われているか」を調査し、人手による評価極性値を算出した。

具体的には、各評価表現が出現する商品レビューに対して、評価表現が文章中で「肯定」と「否定」、あるいは「非評価」のいずれに当てはまるかを判断し、ラベルを付与する。ラベルの付与に際しては、以下の場合に「肯定」または「否定」の評価を判断する。

- ・評価表現が直接的に商品やその要素を評価している場合
  - 例：「この掃除機は素晴らしい」
  - 例：「シャツの生地が薄いのが残念です」
- ・評価表現が商品进行评估する要因となる場合
  - 例：「安い掃除機を探していました」
  - 例：「もう少し濃い味だと嬉しかったです」
- ・評価表現が商品によってもたらされる、人の感情や状態を表す場合
  - 例：「掃除が楽しくなりました」
  - 例：「こんな商品に出会えて幸せです」
  - 例：「少し腰が痛くなります」

一方で、以下のような場合には「非評価」のラベルを付与する。

- ・評価表現が商品とは直接関係のないものを修飾する場合
  - 例：「背の低い私には不向きでした」
  - 例：「こだわりの強い母も喜んでいました」
- ・評価を判断するだけの情報がない場合
  - 例：「この掃除機は小さいです」
  - 例：「普通の生地です」

以上の基準に従い、3種類のラベルを付与した後、次式で人手の評価極性値を決定する。

$$\frac{\text{positive} - \text{negative}}{N} \quad (6.1)$$

ここで、*positive* は「肯定」のラベルを付与した回数、*negative* は「否定」のラベルを付与した回数である。*N* はラベルを付与した回数であり、この実験では  $N = 15$  である。

次に、以上の手順で求めた人手による評価極性値を横軸に、提案手法により決定した評価極性値を縦軸に取ったグラフを図 6.1 から図 6.3 に示す。

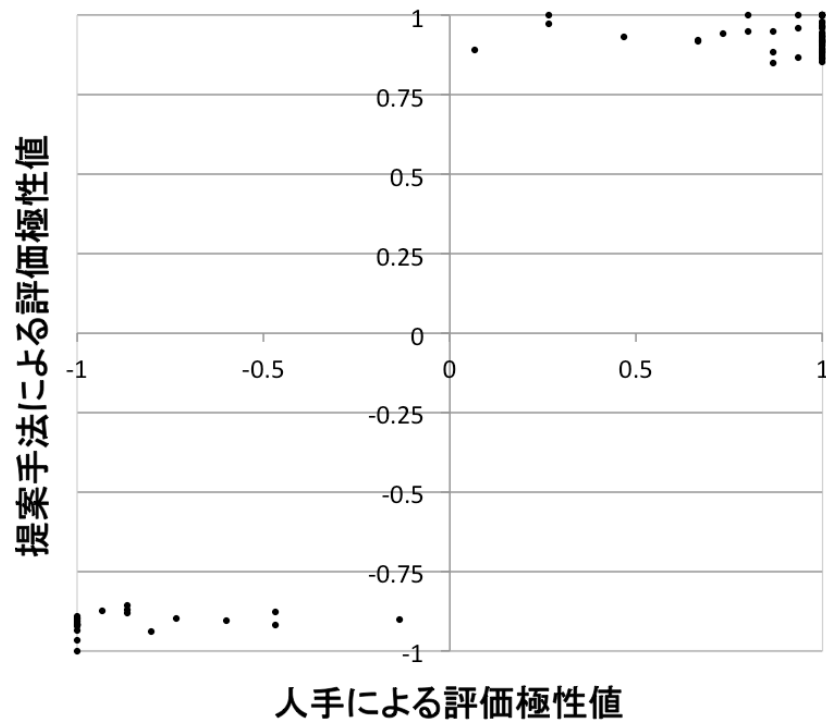


図 6.1 「掃除機」の評価極性値

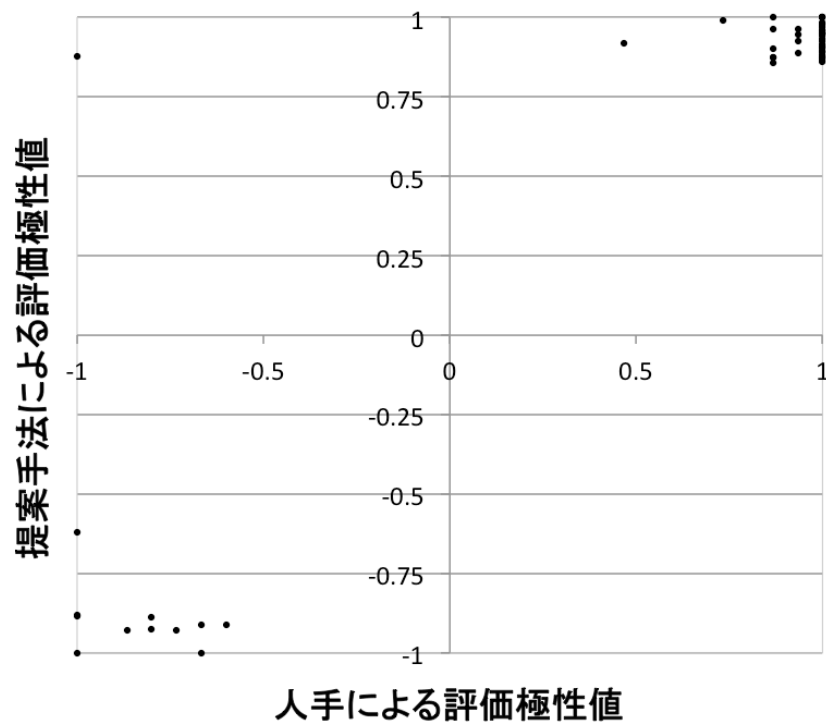


図 6.2 「スーツ」の評価極性値

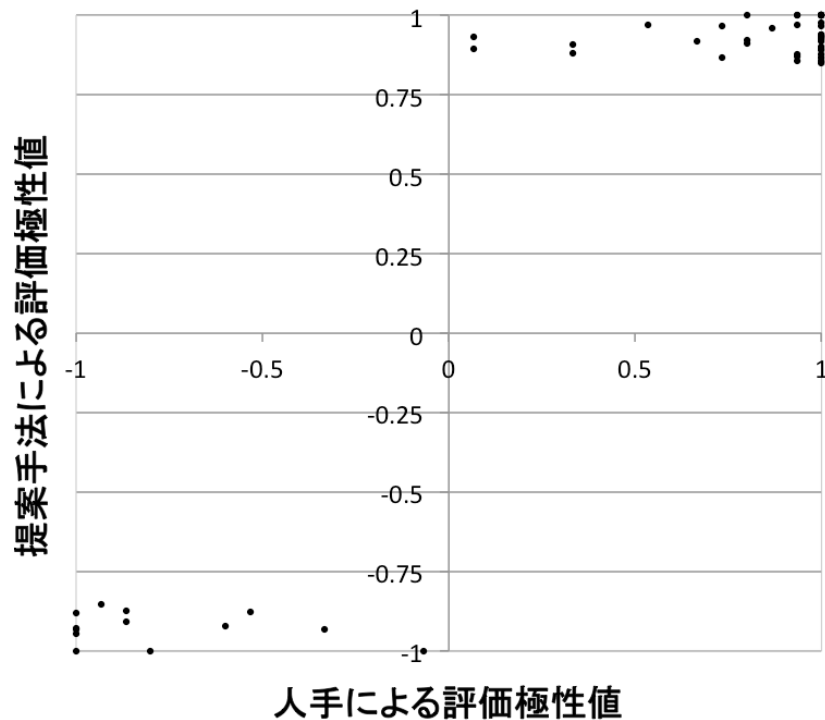


図 6.3 「和菓子」の評価極性値

評価極性値はいずれも正の場合に肯定、負の場合に否定となるため、第一象限と第三象限に点が集中することが好ましい。上記の結果から、いずれのジャンルにおいても高い確率で両者の評価極性が一致していることが分かる。そのため、予め設定した少数の種表現から、商品ジャンル毎の評価表現辞書を構築することができており、提案手法による評価極性値の付与は概ね妥当なものであると考えることができる。

しかし、3つの商品ジャンルのうち、「スーツ」だけは極端に評価極性値が異なるデータ点が存在することが分かる。この評価表現は形容動詞語幹の名詞「駄目」であり、本来なら否定表現として抽出されるのが好ましい語句であるが、今回の実験では肯定表現として抽出されてしまっていた。

そこで、提案システムによって語彙情報の学習時に「駄目」を評価表現候補として抽出した箇所を調査した。その結果、候補として「駄目」を抽出した文章は全部で14箇所存在し、そのうち正しく否定表現として抽出した回数は2回、間違えて肯定表現として抽出した回数は12回であった。また、その12個の商品レビューのうち、内容がほぼ同一のものが複数存在した。商品レビューのフィルタリングは完全に同一内容のものに限るため、一部表現が異なるものは除去の対象とはならない。この重複した商品レビューを考慮から除外すると、誤った抽出箇所は7箇所となった。

次に、この 7 箇所の抽出箇所について、誤った原因をそれぞれ考察した。それらの内訳を表 6.8 に示す。

表 6.8 評価表現候補「駄目」に関する誤った抽出箇所の内訳

原因	抽出数
「価格」に関する評価からの評価極性の誤り	6
条件を表す節からの評価極性の誤り	1

主な原因は表中の 2 つであり、その大半は「価格」に関する評価からの評価極性の誤りであった。このような誤りが発生するレビューの例を以下に示す。

例：「価格が安かったので、品質は駄目だと思っていましたが、大満足でした。」

この場合、「安い」という肯定表現と順接関係にあるため、「駄目」も肯定表現として判定されてしまう。しかし実際には、この文章では「価格が安かった」という肯定の情報が、他の文節には否定的な情報として伝播している。このように、ある要素が良ければ、その引き換えに他方の要素が悪いと想定されるケースは他にも存在する。

例：「小型なのに吸引力もばっちり!」

「掃除機」に関するこの例では、本来肯定的であることが多い「小型」と「吸引力が良い」という情報が、逆接の関係となっている。こういった場合には、本手法による極性反転の処理だけでは対応することができない。基本的にはそのような評価箇所は全体に対して割合が小さく、ノイズとして処理することが出来るが、評価表現候補「駄目」の出現頻度はコーパス中で 98 回と比較的少なく、このような結果になったと考えられる。

次に、条件を表す節からの評価極性の誤りが発生するレビューの例を示す。

例：「タイトすぎるのは駄目なので、L サイズで良かった。」

この場合のレビューも、通常とは異なり肯定表現と否定表現が順接関係となる。後半の文節は「そうではない L サイズで良かった」や「L サイズの方で良かった」といった表現が省略されたものだと考えられ、通常はノイズとして処理される事を期待するものである。いずれの場合にも、対応するには現行の手法に何らかの改善が必要となる。

### 6.2.3 語彙数の評価

6.2.2 節の評価実験では、提案手法が一定の精度で語句の肯定・否定を判定できることを示した。ここでは、提案手法により獲得できる語彙数について評価する。本研究では評価極性を連続値とした手法を採用している。本手法で獲得した評価表現数を、以前の 2 値の評価極性による手法[16]と比較した結果を表 6.9 に示す。

表 6.9 商品ジャンル毎の獲得語彙数

ジャンル名	獲得した語彙数	改善数
掃除機	91	+20
スーツ	59	+6
和菓子	65	+28

この結果から、実験を行った全ての商品ジャンルにおいて、語彙数が改善していることが分かる。これは評価極性を連続値としたことで、評価表現の抽出精度が向上し、ノイズとなる評価表現を増やさずに抽出時の閾値を下げるのが可能になったためだと考えられる。引き続き抽出精度を向上させ、より低い閾値での実験を行うことができれば、さらに多様な評価表現が収集可能だと期待できる。

次に、本研究で抽出できなかった語句について考察する。「掃除機」のジャンルにおいて、種表現から語彙を拡張していく上で抽出することができなかった評価表現候補の出現頻度は図 6.4 のようになった。横軸は語彙数である。

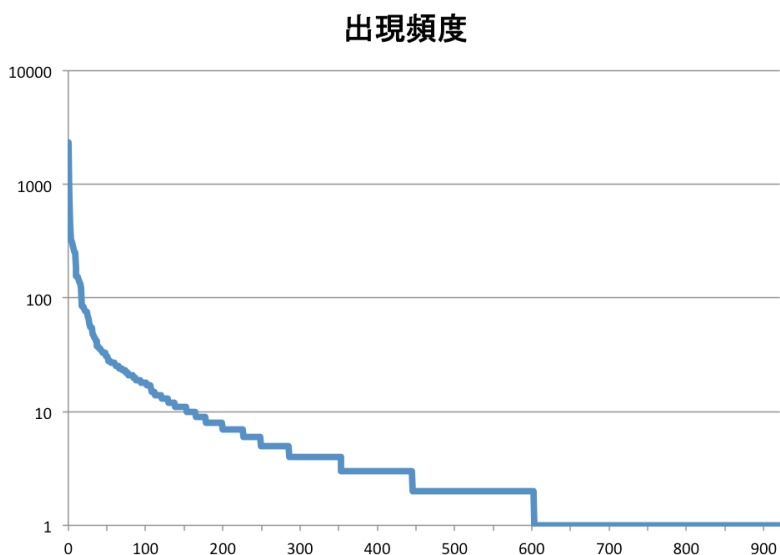


図 6.4 「掃除機」で抽出されなかった評価表現候補の出現頻度

この図から、抽出できなかったものの多くは出現頻度が 10 回に満たないことが分かる。数万件のレビューに対してこの出現頻度は非常に低いため、ここではコーパスに 10 回以上出現した評価表現候補についてのみ言及する。それぞれの商品ジャンルにおいて、抽出可能であった評価表現候補、及び抽出できなかった評価表現候補の数と、抽出可能な候補の割合を表 6.10 に示す。

表 6.10 抽出可能な評価表現候補の割合

ジャンル名	抽出可能	抽出不可	割合
掃除機	518	165	0.758
スーツ	402	185	0.685
和菓子	405	151	0.728

以上の結果，本研究において抽出対象となる評価表現候補のうち，およそ 70%は提案手法によって網羅されていた．特に，学習コーパスにおける出現頻度が 1000 回を超える表現についてはほぼ全て抽出することが可能であった．

しかし，図 6.4 には出現頻度が 100 回を超える評価表現候補も幾つか見られ，「掃除機」の場合は 17 語存在した．それらの語句を表 6.11 に示す．

表 6.11 「掃除機」において抽出できなかった出現頻度上位の評価表現候補

評価表現候補	品詞	出現頻度
にくい	形容詞	2339
づらい	形容詞	797
っぽい	形容詞	430
色々	名詞	322
うまい	形容詞	314
勝手	名詞	293
いろいろ	名詞	271
細か	名詞	256
急	名詞	251
直ぐ	名詞	185
肝心	名詞	156
難しい	形容詞	155
はるか	名詞	153
大量	名詞	146
すごい	形容詞	140
無事	名詞	134
确实	名詞	124



この結果を見ると、抽出できなかった候補の多くは接尾の形容詞と形容動詞語幹の名詞であった。本研究では形容詞に対しての形態素の結合処理は行っていなかった。そのため名詞や形容詞に接続する接尾の形容詞が取り残されてしまったものと考えられる。これを適切に処理すれば、評価表現の抽出精度も向上すると考えられるため、今後取り組んでいく。後者については助詞「に」等と接続することで副詞として機能する語句が多いと考えられる。こちらも同様に形態素の結合処理を改善し、予め副詞として処理することで、収集結果に改善が見られる可能性がある。

### 6.3 特徴語の収集結果

次に、特徴語の収集結果を示す。特徴語候補は評価表現から抽出され、式 4.4 の重要度を付与する。収集した特徴語候補の重要度を降順に並べたものを図 6.5 に示す。

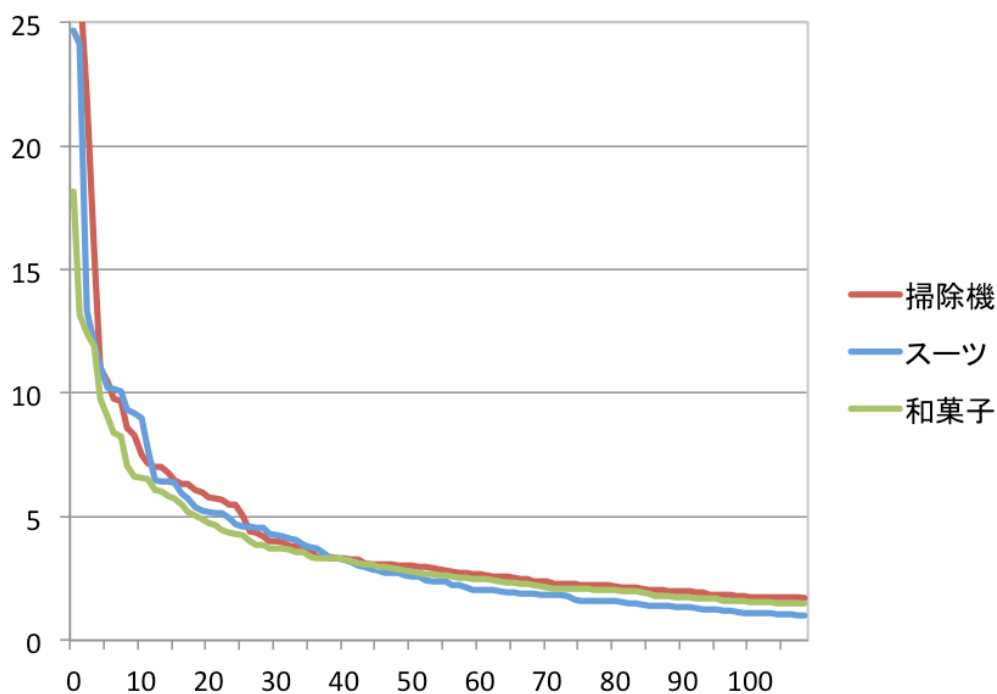


図 6.5 特徴語候補の重要度

この図から分かる通り、特徴語候補の重要度は、4.2.6 節で記述したようなロングテール構造となっている。このうち、累積重要度の上位  $\beta = 50\%$  に含まれるものを特徴語とし、それ以外のロングテール部分は準特徴語とする。特徴語は「評価軸の構築」においてクラスタリングに使われるデータ点となり、準特徴語は特に類似したクラスタがある場合に限り後から分配される。各商品ジャンルにおける特徴語の収集結果を表 6.12 に示す。

表 6.12 各商品ジャンルにおける特徴語の収集結果

商品ジャンル	特徴語
掃除機	音, 使い勝手, 吸引力, 見た目, 値段, 動き, 色, 本体, 評判, お手入れ, 静音性音, お掃除, サイズ, 容量, 価格, ごみ捨て, パワー, 滑り, 組み立て, お値段, ヘッド, 空気, コストパフォーマンス, コード, 性能, 充電時間, 乾き, コードレス, 形, 取り扱い, 動作音, 力, 腰, 難点, レビュー評価, 使い心地, 部屋, 紙パック, モーター音, 調子, 部分, 階段掃除
スーツ	生地, 着心地, 肌触り, 値段, サイズ, シルエット, 質, 品質, コストパフォーマンス, 価格, 評判, 形, 色, 物, 使い勝手, センス, 見た目, 質感, 丈, ライン, お値段, 種類, 乾き, 仕立て, 袖, 着丈, サイズ展開, アイロンがけ
和菓子	味, 評判, 甘さ, 見た目, 値段, あんこ, バランス, 大きさ, 餡, 相性, サイズ, お値段, 皮, 食感, 送料, 香り, お餅, 口当たり, 風味, おもち, 価格, ろう, 形, あん, 量, 大福, パッケージ, クリーム, 縁起, 使い勝手, 抹茶, 外, 色, 単価, 生地, 甘み, 箱, 笹団子, 部分, 彩り, ここの, スイートポテト, 塩加減, ウケ, 腹持ち, 商品, いちご, 餅, 栗, イチゴ, 苺, 抹茶餡, 甘さ加減, 種類, のど越し

「スーツ」のジャンルにおける「物」や、「和菓子」における「ここの」、「商品」は明らかなノイズだが、それ以外は概ね特徴語として機能しうる表現が獲得できていると考える。また、「和菓子」のジャンルに「ろう」という特徴語があるが、コーパス中から「うい」という形容詞が抽出されているため、「ういろう」を誤って形態素解析したものと考えられる。この扱いについては今後、解決策を考えていく必要がある。

## 6.4 評価軸の構築

収集した語彙を用いて評価軸を構築した。始めに、k-medoids 法に使う k の値を決定するために Gap 統計量を用いて最適なクラスタ数を決定した。その結果を表 6.13 に示す。

表 6.13 Gap 統計量を用いたクラスタ数

ジャンル名	クラスタ数
掃除機	14
スーツ	12
和菓子	7

次に、自動決定したクラスタ数を用いたクラスタリング結果を表 6.14 から表 6.16 に示す。評価軸内の最初の特徴語は k-medoids 法における代表点であり、評判分析を行う際、レーダーチャートの軸のラベルとなる。その他の特徴語は代表点との距離の昇順に並んでいる。

表 6.14 「掃除機」のクラスタリング結果

	特徴語群	重要度合計	除去フラグ
評価軸 1	モーター音, 音, 静音性音, 動作音	66.4	F
評価軸 2	使い勝手, 評判, 調子	38.7	F
評価軸 3	吸引力, パワー, 力	32.3	F
評価軸 4	形, 色, 見た目	29.7	F
評価軸 5	価格, 値段, お値段	23.7	F
評価軸 6	動き	10.4	T
評価軸 7	ヘッド, サイズ, 本体, 部分, 難点, 容量, 紙パック	40.0	F
評価軸 8	ごみ捨て, お手入れ, 組み立て, 取り扱い	24.9	F
評価軸 9	使い心地, 滑り, 性能, コードレス, お掃除, 階段掃除	30.2	F
評価軸 10	部屋, 空気	9.09	T
評価軸 11	レビュー評価, コストパフォーマンス	9.50	T
評価軸 12	コード	5.49	T
評価軸 13	乾き, 充電時間	9.43	T
評価軸 14	腰	3.83	T

表 6.15 「スーツ」のクラスタリング結果

	特徴語群	重要度合計	除去フラグ
評価軸 1	生地	24.7	F
評価軸 2	質感, 肌触り, 着心地	43.2	F
評価軸 3	お値段, 価格, 値段	26.4	F
評価軸 4	サイズ展開, 種類, サイズ	20.8	F
評価軸 5	シルエット, ライン, 形, 色	29.7	F
評価軸 6	質, 品質, 見た目, 物, コストパフォーマンス, センス	48.3	F
評価軸 7	評判	8.97	T
評価軸 8	使い勝手	6.41	T
評価軸 9	丈, 袖, 着丈	14.7	F
評価軸 10	乾き	5.12	T
評価軸 11	仕立て	4.95	T
評価軸 12	アイロンがけ	4.55	T

表 6.16 「和菓子」のクラスタリング結果

	特徴語群	重要度合計	除去フラグ
評価軸 1	餡, あんこ, あん, 大福, クリーム, 商品, 風味, 笹団子, スイートポテト, 食感, 抹茶餡, 味, 大きさ, ここの, 量, 相性, ろう, 香り, 口当たり, 種類, 抹茶, のど越し	117	F
評価軸 2	縁起, 使い勝手, 腹持ち, ウケ, 評判, バランス	35.3	F
評価軸 3	甘さ, 甘み, 甘さ加減, 塩加減	21.6	F
評価軸 4	見た目, 形, 箱, パッケージ, 色, 彩り	30.7	F
評価軸 5	お値段, 値段, 価格, 送料, 単価	30.5	F
評価軸 6	いちご, イチゴ, 苺, 栗, サイズ	18.0	F
評価軸 7	生地, 餅, お餅, おもち, 部分, 皮, 外	29.9	F

重要度合計は、評価軸内の特徴語の重要度を合計したものである。この値が小さいものは4.3.3の基準に従って、ノイズの評価軸として除去を行った。除去した評価軸は除去フラグが‘T’となっているクラスタである。以上の結果、それぞれのジャンルについて確信度の高い評価軸のみを残すことができる。

次に準特徴語の中から、特定の評価軸と距離が近いものがあるかを判定し、見つければその評価軸に分配する。各商品ジャンルで分配された準特徴語を表6.17から表6.19に示す。

表 6.17 「掃除機」における準特徴語の分配結果

	代表点	特徴語群
評価軸 1	モーター音	おと, 排気音, 騒音, 若干音, 作動音
評価軸 2	使い勝手	
評価軸 3	吸引力	吸塵力, 吸収力, 威力, 馬力
評価軸 4	形	ピンク, 姿
評価軸 5	価格	金額, その他値段, 料金, 店, 電気代
評価軸 6	ヘッド	ヘッド部分
評価軸 7	ごみ捨て	ごみ処理, 取り出し, 取り付け, 使い方, ゴミ, ゴミ処理, 後片付け
評価軸 8	使い心地	

表 6.18 「スーツ」における準特徴語の分配結果

	代表点	特徴語群
評価軸 1	生地	布地, 記事, 生地自体, 厚み, 秋冬物, パープル, たより生地
評価軸 2	質感	
評価軸 3	お値段	金額
評価軸 4	サイズ展開	品数, バリエーション, 品ぞろえ, カラーバリエーション
評価軸 5	シルエット	状態
評価軸 6	質	
評価軸 7	丈	手, 袖丈, 腕, 衿丈, 股下, 長さ

表 6.19 「和菓子」における準特徴語の分配結果

	代表点	特徴語群
評価軸 1	餡	
評価軸 2	縁起	口コミ, 口溶け
評価軸 3	甘さ	
評価軸 4	見た目	包装紙, 入れ物, 巾着
評価軸 5	お値段	金額
評価軸 6	いちご	
評価軸 7	生地	

準特徴語の分配により, 幾つかの評価軸で語彙数が飛躍的に改善した. それぞれの構築結果についての考察は後述するが, 「金額」を始めとして, 「音」や「吸引力」に関する準特徴語が正しい評価軸に分配されたと考える.

以上の結果, 評価軸が完成する. 最終的な評価軸の構築結果を表 6.20 から表 6.22 に示す. 評価軸内の最初の特徴語は k-medoids 法における代表点であり, その他の特徴語は代表点との距離の昇順に並んでいる.

表 6.20 「掃除機」における評価軸の構築結果

	特徴語群	代表評価表現
評価軸 1	モーター音, 音, おと, 排気音, 静音性音, 動作音, 騒音, 若干音, 作動音	うるさい, 大きい, 静か
評価軸 2	使い勝手, 評判, 調子	良い, よい, いい
評価軸 3	吸引力, 吸塵力, 吸収力, 威力, 馬力, パワー, 力	強い, 十分
評価軸 4	形, 色, 見た目, ピンク, 姿	かわいい, 可愛い
評価軸 5	価格, 値段, お値段, 金額, その他値段, 料金, 店, 電気代	安い, 手頃, 手ごろ
評価軸 6	ヘッド, ヘッド部分, サイズ, 本体, 部分, 難点, 容量, 紙パック	小さい, 軽い
評価軸 7	ごみ捨て, お手入れ, ごみ処理, 取り出し, 取り付け, 組み立て, 使い方, ゴミ, ゴミ処理, 後片付け, 取り扱い	簡単, 楽
評価軸 8	使い心地, 滑り, 性能, コードレス, お掃除, 階段掃除	快適, 満足

表 6.21 「スーツ」における評価軸の構築結果

	特徴語群	代表評価表現
評価軸 1	生地, 布地, 記事, 生地自体, 厚み, 秋冬物, パープル, たより生地	薄い
評価軸 2	質感, 肌触り, 着心地	柔らかい, やわらかい
評価軸 3	お値段, 価格, 値段, 金額	安い, 手頃
評価軸 4	サイズ展開, 品数, バリエーション, 種類, 品ぞろえ, カラーバリエーション, サイズ	豊富
評価軸 5	シルエット, ライン, 形, 状態, 色	きれい, キレイ
評価軸 6	質, 品質, 見た目, 物, コストパフォーマンス, センス	満足, 駄目
評価軸 7	丈, 袖, 手, 袖丈, 腕, 着丈, 衿丈, 股下, 長さ	短い



表 6.22 「和菓子」における評価軸の構築結果

	特徴語群	代表評価表現
評価軸 1	餡, あんこ, あん, 大福, クリーム, 商品, 風味, 笹団子, スイートポテト, 食感, 抹茶餡, 味, 大きさ, ここの, 量, 相性, ろう, 香り, 口当たり, 種類, 抹茶, のど越し	おいしい, 美味しい
評価軸 2	縁起, 使い勝手, 腹持ち, ウケ, ロコミ, 口溶け, 評判, バランス	良い, よい
評価軸 3	甘さ, 甘み, 甘さ加減, 塩加減	程よい, くどい, しつこい, 優しい
評価軸 4	見た目, 包装紙, 形, 箱, パッケージ, 入れ物, 巾着, 色, 彩り	可愛い, かわいい, 綺麗, きれい
評価軸 5	お値段, 値段, 価格, 金額, 送料, 単価	安い, 高い, 手頃, 手ごろ
評価軸 6	いちご, イチゴ, 苺, 栗, サイズ	大きい, みずみずしい, 甘酸っぱい
評価軸 7	生地, 餅, お餅, おもち, 部分, 皮, 外	やわらかい, 柔らかい, カリカリ

以上の結果より、いずれの商品ジャンルにおいても、商品レビューから複数の評価軸を構築することが出来ていることが分かる。通常のレビューサイトでは、商品の評価は1次元の尺度でしか行われていない。幾つかのネットショッピングサイトでは商品に対して評価軸が設定されているが、それらは予め設定されたものに対してユーザが点数を付けているため、柔軟性に欠ける。提案手法では商品レビューから評価軸の自動構築を行うことで、ユーザが興味のある項目を評価軸として扱うことができ、既存のものよりも有効であると考えられる。

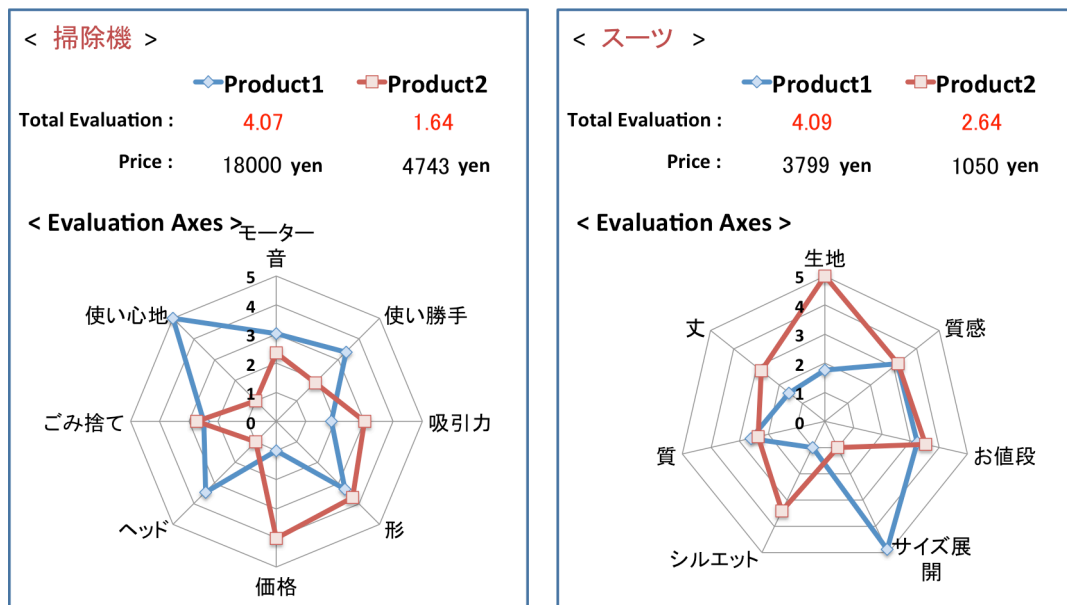
構築結果を詳しく見ると、いずれのジャンルからも「価格」に関する評価軸が構築されていることが分かる。これはどんな商品にも共通な評価軸だと考えられ、商品ジャンルによらず抽出できた事は好ましい結果だと言える。一方で、それぞれの商品ジャンルに特有な評価軸も構築することができた。「掃除機」のジャンルでは「吸引力」、「音」、「本体」、「ごみ捨て」等の評価軸が、「スーツ」のジャンルでは「生地」、「質感」、「バリエーション」、「丈」に関する評価軸、「和菓子」のジャンルでは「甘さ」「外見・包装」「皮」「果物」といった評価軸が構築されていると思われる。「和菓子」の評価軸1と評価軸2は一見すると上手く構築が出来なかったように見受けられる。しかし代表的な評価表現から類推すると、評価軸1は「餡」や商品そのものについての味を表現しているとも考えられる。しかし、味とは直接関係のない特徴語も幾つか分類されてしまっているため、改善の余地がある。また、評価軸6は「具」としてのまとまりは良いが、語彙数が少ないため評判分析には不利であるほか、レーダーチャートのラベルが「いちご」になってしまうという問題がある。

以上より、評価軸の構築という課題においては、「掃除機」や「スーツ」に比べて「和菓子」のジャンルは困難だという結論に至った。前者2つは機能や外見に個体差が少なく、ユーザが扱う語句の違いも少ない。一方で、「和菓子」は形状や食感、材料等の要素が商品毎に大きく異なるため、クラスタリングが困難になるようだ。

## 6.5 商品の評判分析

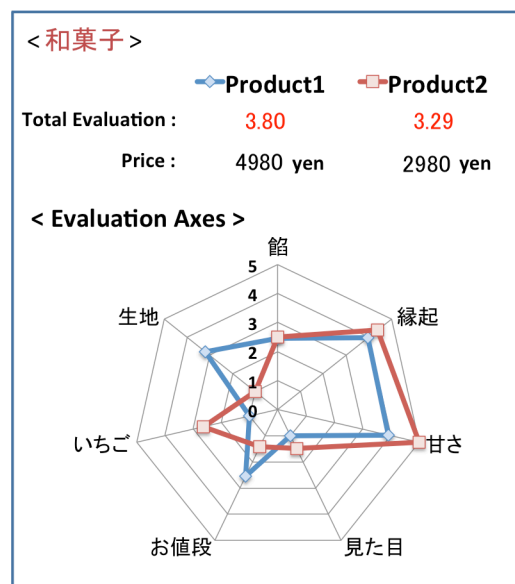
### 6.5.1 レーダーチャートによる可視化

構築した評価軸と評価表現辞書を活用し、それぞれの商品ジャンルにおいて 2 商品を対象に評判分析実験を行った。5 章で述べた「総合評価 (Total Score)」と「評価軸ごとの評価」を各商品に対して分析した結果を図 6.6 に示す。なお評価値はいずれも [1,5] の範囲で正規化を行っており、通常のレビューと同様に、値が高いほど良い評価である。



(a) 掃除機

(b) スーツ



(c) 和菓子

図 6.6 商品ジャンル毎の評判分析結果

これにより、複数の評価項目の中から自身の趣向に合わせて商品を選択することが可能となる。「掃除機」の例では、「使い心地」や「ヘッド」、「モーター音」といった使用感で優位な商品（Product1）と、「価格」に対する評価が極めて高い商品（Product2）を比較して選択することができる。

以上のように、評価軸と評価表現辞書により、多面的な評判分析が可能となった。楽天市場での商品レビューは通常 1 から 5 の評価値でしか表現されていないが、このシステムを利用すれば特定の評価項目についての評価を簡単に知ることができ、他の商品との比較が容易になると期待される。

## 第 7 章 おわりに

### 7.1 まとめ

本研究では、ネットショッピングサイトにおける商品選択を支援する技術として、新たに評価軸を定義した評判分析手法を提案した。評価実験では 3 つの商品ジャンルを対象として評価軸ごとの評判分析を行い、提案手法が幅広い商品ジャンルに対して適用可能であることを示した。

評判分析の多くの関連研究では、レビューを肯定と否定という 1 次元の尺度でしか分析することができなかったが、提案手法では商品ジャンル毎の評価軸に基づいて評価を分析することが可能である。また、多次元の評判分析や評価軸情報の構築では、複数の評価項目を判別するための種語集合が必要である場合が多く存在するが、本研究において必要なのは少数の種表現だけである。この種表現は商品ジャンルに依存せず、肯定または否定を示すものを登録するため、学習コーパスを変更するだけでその商品ジャンルに適した評価軸を構築することが可能である。既存研究では商品ジャンルを変更する際に種語集合も変更しなければならず、その点において本研究は汎用性が高いと考えられる。

構築した評価表現辞書に関しては、評価極性値の妥当性と獲得した語彙数の両方について考察を行い、本研究が有効であることを示した。評価表現の評価極性を連続値としたことで、肯定語や否定語としての確信度が高いものに強い影響力を持たせることが可能となるため、評判分析の精度向上が見込める。また、評価極性を 2 値とした場合よりも、評価極性値の閾値を下げた場合の抽出精度が落ちにくいと考えられるため、平均的に以前よりも収集できる語彙数が多くなった。評価軸情報については、正解データが存在しないため定量的な評価には限界があるが、主観による評価において一定の成果を得られたと考える。以前の我々の研究では、クラスタリングに用いる  $k$  の値は予め指定する必要があり、しばしば類似した特徴語が別の評価軸に分類される事があった。Gap 統計量を用いて  $k$  の値を決定することで人手による決定の必要がなくなったほか、クラスタひとつひとつの凝集性が向上し、ノイズが減ったように思われる。

## 7.2 今後の課題

今後の課題は、語彙情報の収集結果の改善である。評価極性値を連続値としたことで獲得した語彙数は向上したが、依然として抽出できない評価情報は多く存在する。現状では、評価表現候補として扱うのは形容詞と形容動詞語幹の名詞のみであるため、「吸引力がある」といった動詞による評価情報は抽出することができない。しかし、評価表現候補に動詞を加えた場合、抽出される候補が爆発的に増え、評価表現全体の抽出精度が下がってしまう可能性がある。これについては、既存の幾つかの研究を手掛かりに、「複合表現を評価表現として扱う」ことにより、抽出が可能だと考えられる。「特徴語候補と評価表現候補が係り受け関係にある」といった条件を満たした場合に、それらの共起ペアを評価表現の候補とする。こうすることで動詞をノイズとすることなく、必要な表現だけを辞書に登録できると考える。加えて、評価情報が抽出できない商品レビューとして、「この味で、この価格は安いです」というものが存在した。この例では「味」と係り受けする肯定表現がなく、「味」についての肯定評価を抽出することが出来なかった。このような例については、係り受けによって評価された特徴語にも評価極性値を持たせることで、評価極性値を伝搬させることで解決できると考えられる。この実装についても今後取り組んでいく。

また、6.2.3 節で記述したとおり、形態素同士の処理を改善する必要があると考えられる。抽出されなかった形容詞の中には接尾のものが複数存在し、それらは前の形態素と結合することで、別の形容詞として抽出可能であったと考えられる。今回、その解決には至らなかったが、システムの精度が改善されると期待できるため、今後取り組んでいく。加えて、全ての形容詞を評価表現候補とすることが妥当であるか等、品詞の扱いについて一度検討し直す必要性を感じた。

本手法では、語彙情報を判定するために設定する閾値が多数存在する。今回の実験では商品ジャンルを 3 つ採用し、それらについて安定して実験結果が得られる値を探したが、この値が本当にどの商品ジャンルに対しても最適であるかどうかには疑問が残る。加えて「評価極性値」や「特徴語の重要度」の算出方法が適切かどうかについても検討が十分であるとは言えない。評価軸の構築で「和菓子」のジャンルが比較的困難であったように、対象とする商品ジャンルによって評価される特徴語は異なり、文章の構造や語句の種類にも差異が存在する。今回実験を行った商品ジャンル以外も対象とすることで、提案システムが得意とする分野、不得意な分野を知り、さらなるシステムの改善に取り組むことが重要だと考える。

## 謝辞

本研究の実験では、楽天株式会社と国立情報学研究所が公開しているデータセットである「みんなのレビュー・口コミ情報」を使用させていただきました。心からお礼を申し上げます。

また、本研究を進めるにあたって、ご指導していただいた沼尾雅之教授に深く感謝を申し上げます。同じく、日頃から意見やアドバイスを頂いた沼尾研究室の皆さまにも感謝いたします。

## 参考文献

- [1] 稲葉真純, 長野伸一, 長健太, 溝口祐美子, 川村隆浩, “CGM 分析技術の現状と課題,” 人工知能学会研究会資料, SIG-SWO-A603-06, 2007.
- [2] 林田英雄, “Web マーケティングのための CGM 分析,” Unisys 技報 : Unisys technology review. 31(3)(110), pp.275-285, 2011.
- [3] 乾孝司, 奥村学, “テキストを対象とした評価情報の分析に関する研究動向,” 自然言語処理 Vol.13, Num.3, pp.201-241, 2006.
- [4] 佐野大樹, “日本語アプレイザル評価表現辞書 (JAppraisal 辞書) ~態度評価編~ Version1.2 仕様説明書, 及び, 評価表現分類表,” 言語資源協会, 2012.
- [5] 小林のぞみ, 乾健太郎, 松本裕治, “意見情報の抽出/構造化のタスク仕様に関する考察,” 情報処理学会研究報告 NL-171, pp.111-118, 2006.
- [6] Kamps, J., Marx, M., Mokken, R. J., and de Rijke, M. "Using WordNet to Measure Semantic Orientations of Adjectives," Proc. of the Fourth International Conference on Language Resources and Evaluation (LREC-2004) Vol.4, pp.1115-1118, 2004.
- [7] 那須川哲哉, 金山博, “文脈一貫性を利用した極性付評価表現の語彙獲得,” 情報処理学会自然言語処理研究会(NL-162-16), pp.109-116, 2004.
- [8] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一, “意見抽出のための評価表現の収集,” 自然言語処理 Vol.12, No.2, pp.203-222, 2005.
- [9] グェン ファム タン タオ, 岡部誠, 尾内理紀夫, 林貴宏, 西岡悠平, 竹中孝真, 森正弥, “新たな弱教師付き型分類手法 Bautext”, 情報処理学会論文誌 Vol.52 No.1, pp.269-283, 2011.
- [10] 中野裕介, 湯本高行, 新居学, 上浦尚武, “機械学習による商品レビューの属性-意見ペアの抽出,” 研究報告データベースシステム(DBS) Vol.2015-DBS-162, No.14, pp.1-8, .2015.
- [11] 高野敦子, 池奥渉太, 北村泰彦, “因果関係に着目した口コミ Web サイトからの評価表現抽出,” 人工知能学会論文誌 Vol.24, No.3, pp.322-332, 2009.
- [12] Peter D. Turney, "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews," Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), pp.417-424, 2002.



- [13] 鈴木泰裕, 高村大也, 奥村学, “Weblog を対象とした評価表現抽出,” 人工知能学会セマンティックウェブとオントロジー研究会 SIG-SW&ONT-A401-02, 2004.
- [14] 熊本忠彦, 河合由起子, 田中克己, “新聞記事を対象とするテキスト印象マイニング手法の設計と評価,” 電子情報通信学会論文誌 D Vol.J94-D, No.3, pp.540-548, 2011.
- [15] 前川浩基, 中原孝信, 岡田克彦, 羽室行信, “大規模ニュース記事からの極性付き評価表現の抽出と株価収益率の予測,” オペレーションズ・リサーチ : 経営の科学 Vol.58, Num.5, pp.281-288, 2013.
- [16] 金兵裕太, 沼尾雅之, “ネットショッピングサイトの商品レビューを利用したジャンル毎の評価軸の自動構築,” The 8th Forum on DataEngineering and Information Management, C2-3, 2016.
- [17] 杉浦広和, “議事録集合からの特徴語抽出とその応用に関する研究,” 名古屋大学電気電子・情報工学科卒業論文, 未公刊, 2009.
- [18] Dan Pelleg, Andrew W. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters," ICML '00 Proceedings of the Seventeenth International Conference on Machine Learning, pp.727-734, 2000.
- [19] 石岡 恒憲, “クラスター数を自動決定する k-means アルゴリズムの拡張について,” 応用統計学 Vol.29, No.3, pp.141-149, 2001.
- [20] Robert Tibshirani, Guenther Walther and Trevor Hastie, "Estimating the number of clusters in a data set via the gap statistic," J. R. Statist. Soc. B Vol.63, Issue 2, pp.411-423, 2001.

## 付録 評判分析システム

付録では、第 4 章および第 5 章で提案した、語彙情報の学習とレーダーチャート生成の機能を実装した評判分析システムを紹介する。本システムは **Java** によって実装を行っており、係り受け解析は **CaboCha**<sup>1</sup>を用いている。また、クラスタリング等の一部の処理では、**R** も利用している。

本システムは、**CaboCha** による係り受け解析を実行するプログラム、評価表現辞書と評価軸を実装し、学習と分析等の処理機能を備えたプログラム、そしてその処理機能を GUI から呼び出すプログラムで構成されている。ここでは、GUI によるシステムの使用例を示していく。

はじめに、本システムの起動直後の画面を図 A-1 に示す。



図 A-1 システム起動直後の画面

<sup>1</sup> <https://taku910.github.io/cabochoa/>

次に、扱う商品ジャンル辞書を指定する。新しく辞書を作成する場合には、メニューの [ファイル] から [新規作成] を選択し、「新規ジャンル名」や「学習コーパス」、「種表現」といった必要項目を入力する。また、本システムでは学習後の辞書データを MySQL<sup>2</sup>に保存することが出来る。作成済みの辞書を使う場合は [ファイル] から [開く] を選択する。ここでは、「掃除機」のジャンルを例として取り上げる。図 A-2 は新しく辞書を作成する際のダイアログである。



図 A-2 辞書作成時のダイアログ

ジャンル辞書を開くと、ウィンドウ左側の「Information Panel」と右上の「ジャンル辞書の構築情報」の各項目に情報がセットされる。この時点では評価表現は種表現として設定したものだけであるが、メニューの [辞書] から [語彙の獲得] を選択することで、現在の評価表現を用いて語彙を拡張することができる。種表現を用いて 1 度だけ語彙の獲得を行った際の画面を図 A-3 に、結果が収束するまで語彙を獲得した結果を図 A-4 に示す。

なお、本手法では学習コーパスを複数回にわたって走査するが、その度に係り受け解析を行うことは非効率である。そのため、予め係り受け解析を行った結果をファイル出力しておくことで、処理時間の短縮を行っている。

<sup>2</sup> <https://www-jp.mysql.com>



図 A-3 1 度目の語彙の収集結果



図 A-4 最終的な語彙の収集結果

語彙の収集が完了したら、メニューの「辞書」から「評価軸の構築」を選択する。評価軸の構築が完了すると、図 A-5 で示すように評価軸情報がパネルに表示される。



図 A-5 評価軸の構築結果

以上の手順で、第 4 章で記述した「語彙情報の学習」が終了する。実際の運用にあたっては、語彙情報の学習は開発者（企業）が予め行っておくことを想定している。

続いて、構築した語彙情報を用いた評判分析について説明する。今まで辞書の情報を表示していたパネルは「Dictionary」タブの内容であった。「Analysis」タブを選択することで、評判分析に関するパネルが表示される。そのパネルから「商品選択」ボタンを押すことで分析対象となる商品を選択でき、「評判分析」ボタンで分析が開始する。「商品選択」ボタンによって表示される商品選択用のダイアログを図 A-6 に、「評判分析」ボタンで表示される商品の分析結果を図 A-7 に示す。なお、図 A-7 のレーダーチャートの表示には、フリーのグラフィブラリである JFreeChart<sup>3</sup>を利用している。

<sup>3</sup> <http://www.jfree.org/jfreechart/>

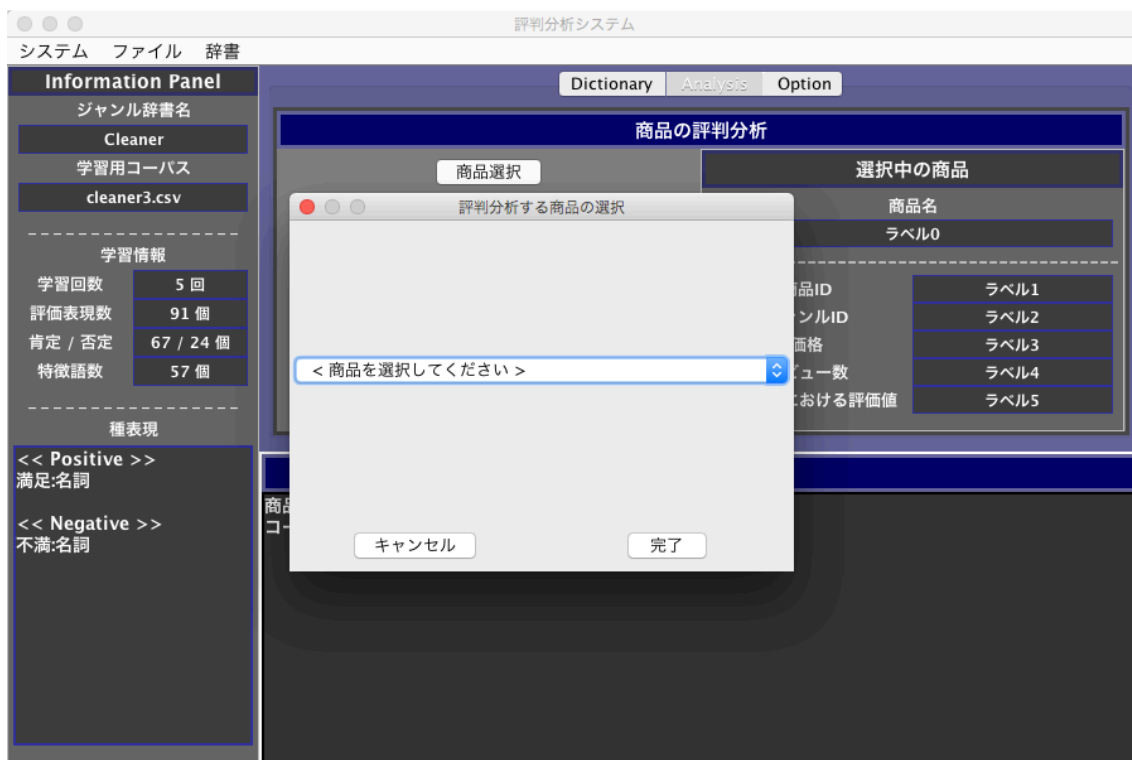


図 A-6 商品選択用のダイアログ

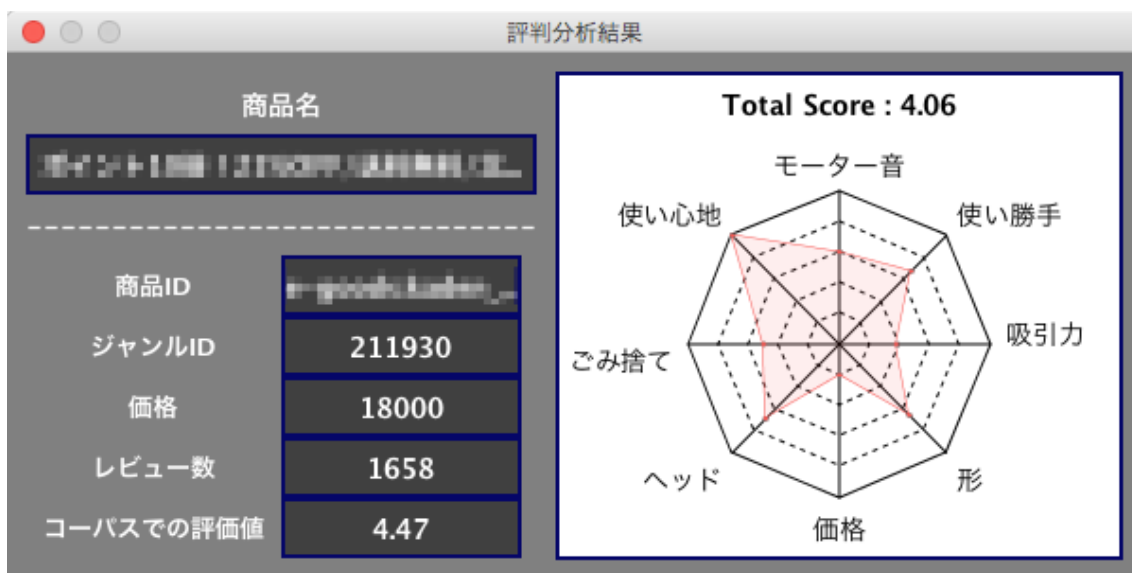


図 A-7 商品の評判分析結果のダイアログ

このように、本システムでは語彙情報の学習から評判分析結果の可視化までを一貫して行うことが可能である。現状は楽天市場のコーパスに対応した実装となっており、多様な形式のコーパスに対応可能なインターフェースは開発中であるが、一連の処理結果を確認するツールとしては十分に機能すると考える。