

評価視点別の言及度を用いた意見文の分類手法の提案

新井 智也[†] 佐藤 哲司^{††}

[†] 筑波大学図書館情報専門学群 〒305-8550 茨城県つくば市春日 1-2

^{††} 筑波大学大学院図書館情報メディア研究科 〒305-8550 茨城県つくば市春日 1-2

E-mail: [†]s0813158@u.tsukuba.ac.jp, ^{††}satoh@slis.tsukuba.ac.jp

あらまし ユーザの商品購入や企業のマーケティングに役立つ Web 上の評判情報が増え続けている．増え続ける膨大な評判情報の中から，ユーザが必要とする情報を得るには多大な時間を要し，効率的な探索方法が必要とされている．本論文では，評判情報が記述された個々の意見文を，言及している評価視点に基づいて分類する意見文分類手法を提案する．評価視点と共起する語に関連度を与え，評価視点ごとに関連語辞書を作成することで，意見文に出現する関連語から各評価視点に対する言及度を算出する．これにより，評価視点の語を含まない意見文も分類可能とする．インターネット上で実際に運用されているレビューサイトの記事を用いた評価を行い，提案法の有効性を確認したので報告する．

キーワード 評判情報，分類，情報検索，評価視点，言及度

Classification Method for Reviews using Degree of Mentioning each Viewpoint

Tomoya ARAI[†] and Tetsuji SATOH^{††}

[†] School of Library and Information Science, University of Tsukuba

1-2, Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

^{††} Graduate School of Library, Information and Media Studies, University of Tsukuba

1-2, Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

E-mail: [†]s0813158@u.tsukuba.ac.jp, ^{††}satoh@slis.tsukuba.ac.jp

Abstract Online reputation is the quite useful information for both product choice and marketing research. The volume of reputation information is continuously increasing. As a result, the user must seek much reputation information to get the useful information. This is a burden very much for a user. In this paper, we propose an approach to classify opinion sentences in either of plural evaluation viewpoints. In this approach, we make related terms dictionary with a related degree for each evaluation viewpoint and we calculate a reference degree every evaluation viewpoint of a sentence with the dictionary. We realize the classification of the opinion sentence with the reference degree. Through comparing proposed approach with SVM, we confirm that precision improved by the proposed approach.

Key words Reputation information, classification, Information retrieval, Evaluation viewpoint, Reference degree

1. はじめに

近年 web 上で個人が情報を発信できる環境が整ってきている．特に，商品やサービスの利用者が投稿する使用感や意見などの評判情報は，商品やサービスの購入を考えているユーザに対して有益な情報であり，積極的に投稿されている．価

格.com^(注1)や楽天トラベル^(注2)など評判情報をレビューやクチコミとして扱うサイトも現れ，ここに書かれているレビューの数も増加してきている．しかし，商品やサービスに対する評判情報が大量に蓄積されていくとともなって，評判情報を活用しようとするユーザにとっては，すべてに目を通すことが困難に

(注1): <http://kakaku.com/>

(注2): <http://travel.rakuten.co.jp/>

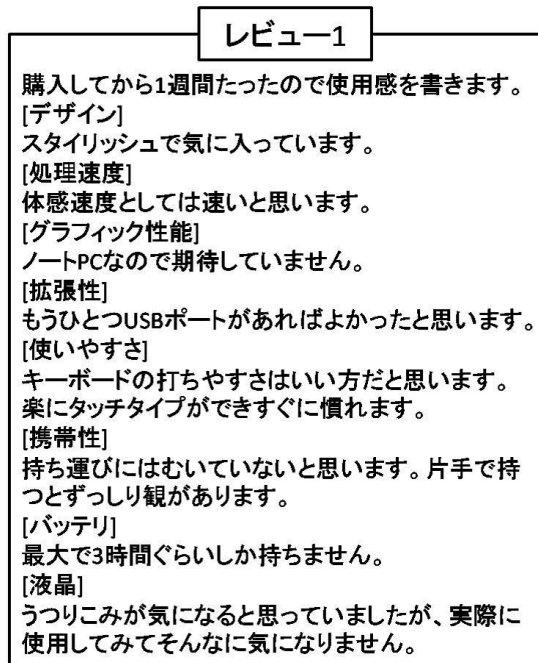
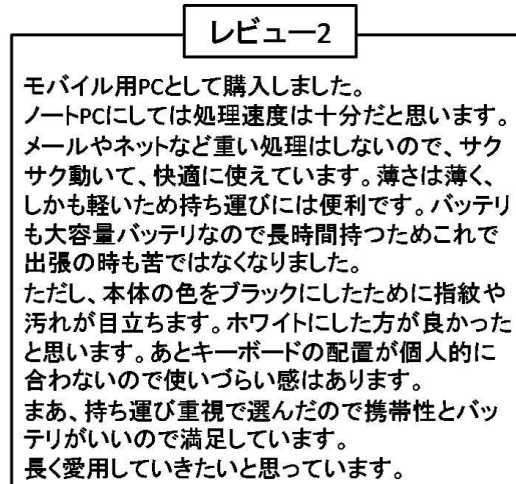


図 1 レビュー記事の例



なり、必要とする情報を見つけるのに多くの時間が必要となっている。

価格.com^(注3)に投稿されたノート PC のレビューを参考に、レビュー記事の一般的な書式を図 1 に示す。図 1 のレビュー 1 に示すように、レビューは複数の視点から評判情報が書かれていることが多い。図 1 の例では、視点として「デザイン」「処理速度」「グラフィック性能」「拡張性」「使いやすさ」「携帯性」「バッテリー」「液晶」の 8 方向から書かれている。この例のようにレビューは、対象となる商品全般について書かれるだけでなく、商品の側面となる様々な視点から書かれていることが多い。レビューやクチコミの中で言及されている、商品やサービスの様々な側面である視点を評価視点と定義する。例えば、あるノート PC に対するレビューに「デザインがとても気に入りました。」という文があった場合、ノート PC の「デザイン」について言及されていることから、この文の評価視点は「デザイン」となる。

評判情報は自由に表記できることから、商品全般について書かれている場合も多い。しかし、図 1 のレビュー 2 に見られるように評価視点にそって明確に文が分けて書かれていない評判情報でも、文単位で見えていくと、評価視点ごとに評価が書かれていることが多い。

以上のことから、評判情報はさまざまな評価視点から評価が書かれていることが多く、評価視点ごとに分類することで、ユーザが閲覧したい評判視点の評判情報のみを閲覧可能にし、ユーザの負担を軽減することが期待できる。

本論文では、評判情報を扱う粒度をレビュー記事単位でなくレビューの中に書かれた個々の意見文単位とし、ユーザが指定した評価視点別に意見文进行分类することを目的とする。意見文

とは「いい」「素晴らしい」などの評価表現を含む意見のみではなく、評価に値する事実文も含める。例えば「デザインがいい」は評価を含む意見文、「バッテリーが 3 時間持ちます」という文は、評価は記述されていないが評価に相当する事実が記述された意見文である。ユーザが指定する評価視点とは、ユーザが閲覧したいと考えた評価視点で、本論文ではユーザが与えるものとする。また、評価視点の語を直接含まない文でも、その評価視点に対する意見文であれば分類できるようにする。この評価視点の語を直接含んでいない文を意見文の中でも特に関連文と呼ぶ。

以下、2 章で関連研究を紹介し本研究の位置づけを明らかにする。3 章で本研究の提案する手法について説明する。4 章で提案する手法について実装を行う。5 章で評価を行い、6 章で考察する。7 章で本研究の総括を行う。

2. 関連研究

評判情報を扱う研究は、主にテキスト評価分析の要素技術に関する研究と、テキスト評価分析の応用研究に大別される [1]。前者は、評価表現辞書の構築に関する研究と、評価情報を観点とした文書または文の分類、評価情報の要素組の抽出に分類される。以下では、本研究に特に関連のある評価表現を用いた分類に関する研究とテキスト評価分析の応用研究の一つである評価分析システムに関する研究について述べる。

2.1 評価表現を用いた分類に関する研究

評価表現を用いた分類が数多くなされているが、分類の単位として文を用いる研究は、Hu [2] ら、箆島ら [3] がある。

Hu らは、種表現をもとに語彙ネットワークを利用して評価表現辞書作成し、その辞書を使って文の中に現れる評価表現の比率をもとに分類をしている。箆島らは、特に文の特徴を考慮して、文内の語の系列情報を利用した方法で分類をしている。

(注 3): <http://review.kakaku.com/>

2.2 評価分析システムに関する研究

テキスト評価分析の応用研究は、評価分析システムに関する研究が盛んである。ユーザにとって負担のない情報量で、提示することを狙いとして、分析データの集約機能や、分析結果の可視化機能を高度化する研究が多い。

立石ら [4] は、レーダーチャート形式で意見を可視化している。ブーストラップ手法を用いて属性表現辞書を作成する際に収集した属性に対して、着眼点（本研究では分類の指標とする評価視点）のラベルを付与し、そのラベルをもとに分類している。辞書に登録する属性は人手で判断し、ラベルの付与も人手で行っている。

赤木ら [5] は、文書がそれぞれの評価属性語に言及している程度を尺度化する手法を提案している。ページ集合の中から出現回数の多い形容詞を利用し、その隣接にある名詞句を使って評価属性語（評価視点）を抽出する。言及度は共起度と文書内の出現回数から算出するが、言及度を使った文書分類までは行っていない。

瀬藤ら [6] は、評価視点ごとに評判情報を分類するシステムを提案している。評価視点の抽出に際して商品説明ページを用いて抽出の精度を高めている。

谷本ら [7] は、特定の商品について述べたテキストを評価属性ごとに点数化している。評価表現辞書として、評価値辞書、評価属性辞書、評価副詞辞書の3種類の辞書を用意しスコアを算出している。本研究の評価視点と対応する評価属性の辞書の作成では、主属性（本研究では分類の指標とする評価視点）に、類似表現や類語を用いて関連属性を抽出し、主属性に結び付けている。

2.3 本研究の位置づけ

本論文では立石らと同様に分類の指標とする評価視点別に分類する手法を提案する。ただし、評価視点抽出は必要としない。また、赤木らと同様に評価視点に対する言及度を求めるが、関連語ごとに関連度を持った関連語辞書を作成し、関連語辞書を用いて文の言及度を求めている点で異なる。ここで言及度とは、文が各評価視点に対してどの程度詳しく言及しているかを表したスコアである。関連文も含めて意見文を分類するという点でも従来研究とは大きく異なる。

3. 意見文分類手法の提案

3.1 提案手法の概要

評価視点別に意見文を分類するにあたり、最初に用語を定義する。

評価対象 評判情報の対象となっている特定の商品やサービス

評価視点 評価対象の性質や特徴または評価対象の一部

関連語 評価視点と関係のある語

関連度 関連語が評価視点とどの程度関係があるかを表したスコア

意見文 評価視点に対して言及している文

言及度 意見文が評価視点に対して言及している詳細度

関連文 評価視点を直接含まない意見文

評価視点に基づいて意見文を分類する単純な手法は、評価視

点の語を含む文を抽出し分類する手法が考えられる。しかし、この手法では評価視点の語を含まない関連文を抽出することはできない。例えば、「デザインが気に入っています。」や「形はスタイリッシュでかっこいいです。」という意見文は、いずれもデザインに対して言及していると考えられるが、後者はデザインという語含んでいないため、単純な手法ではデザインに言及しているとはならない。

そこで本論文で、同じ評価視点を持つドメイン（ノートPCなど）のレビュー集合をコーパスとして、評価視点ごとに関連語辞書を作成する。作成した関連語辞書を用いて各評価視点に対する文の言及度を算出し、関連文も抽出できるようにする。文よりも大きい粒度、例えば段落やレビュー記事を単位として評判情報を扱った場合は、複数の評価視点に対する評判が混在することから、評価視点ごとの分類が行えないため、本論文では文単位として評判情報を分類する。

評判情報（レビューテキスト）を文単位に分割する手順を以下に示す。

- (1) 「。」,「」,「」,改行が2つ以上連続した場合に分割
- (2) (1)で分割されたエレメントの中で、文末を意味しない「、」,「」,「」に接続して現れる改行を除去
- (3) (2)の処理を行ったエレメント中の改行の数が3つ以上ある場合は、改行でそのエレメントを分割し、分割されたエレメントを新たなエレメントとする

この手順で分割されたエレメントを文として扱う(3)を行った理由は、評判情報は統制された表記法で表記されていないため、文末に「。」,「」,「」などの記号を用いず、代わりに改行を用いて文の区切りを表しているレビューも多く見られたからである(3)の前に(2)を行っている理由は「、」,「」,「」に接続して現れる改行は、文末を意味しない改行であると考えられることができるためである。

提案手法の概要を図2に示す。入力は、同じ評価視点を持つドメインのレビュー、分類の対象となるレビュー、複数の評価視点とする。まず、同じ評価視点を持つドメインのレビューと評価視点から、評価視点ごとに関連語辞書を作成する。次に、関連語辞書を使って文の各評価視点における言及度を算出する。最後に、言及度をもとに意見文を分類する。以下、この順で詳細に説明する。

3.2 関連語辞書の構築

3.2.1 関連語の抽出

関連語ごとに関連度を与えた関連語辞書を作成する。関連語に用いる素性は、品詞が名詞と形容詞のものとした。この素性を用いた理由は、文がどの評価視点に対する意見文であるかを特定するために、この素性が有効であると考えられるためである。例えば、「処理がとても速いです。」という意見文では、名詞は「処理」で、形容詞は「速い」である。どちらも評価視点「処理速度」に言及している文に頻繁に現れる語であるため、これらの語は「処理速度」に言及している意見文であるということを決定づけるのに有効な語であると考えられる。文を形態素

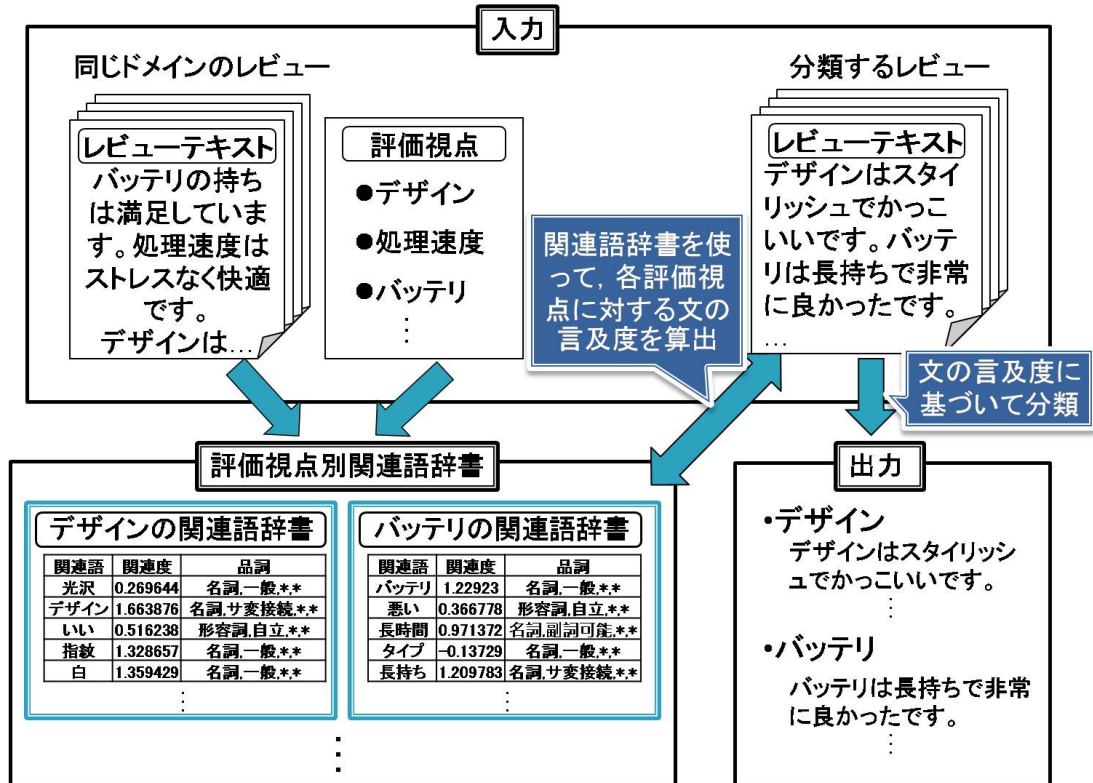


図 2 提案手法の概要

解析器 MeCab^(注4)に入力し、出力結果からこの素性をもつ形態素を関連語とした。

ただし、経験的に関連語にはなりにくい「こと」(< 名詞, 非自立, 一般,*>)や「にくい」(< 形容詞, 非自立,*>)のように品詞に< 非自立 >を含むものと、< 名詞, 一般 >の漢字 1 文字以外の 1 文字は抽出しない。

3.2.2 関連度

関連度の算出の流れは、まず、レビューを分割した文の中から、評価視点の語を含む文を抽出する。次に、レビュー全体における関連語の出現確率と、評価視点の語を含む文の集合における関連語の出現確率(評価視点の語と共起する確率)を求める。最後に、求めた 2 つの出現確率を使って関連度を求める。

関連度の計算式を次式に示す。

$$R_{w(v-j)} = \log \left(\frac{|S_{v-j} \cap S_w|}{|S_{v-j}|} \times \frac{1}{\frac{\sum_{i=1}^n |S_{v-i} \cap S_w|}{\sum_{i=1}^n |S_{v-i}|}} \right) \quad (1)$$

$R_{w(v-j)}$ は評価視点 $v-j$ における関連語 w の関連度を表す。 S_{v-j} , S_w はそれぞれ評価視点 $v-j$ の語を含む文の集合、関連語 w を含む文の集合を表している。変数 i は $1 \leq i \leq n$ の値で、 n は分類の指標とする評価視点の数である。関連度は、評価視点 $v-j$ と関連語 w が共起する確率と、文の全体の集合における関連語 w の出現確率の積を対数とした値とする。文の全体の集合は、各評価視点の語を含む文の集合をすべての評価視点

について合わせた集合である。これは、意見文がどの評価視点に対して言及しているのかを特定しやすくするために、その評価視点のみに頻繁に現れる語に、高いスコアの関連度を持たせるためである。

関連度は以下の理由から対数で表すこととした。評価視点の語と共起する確率が小さい語でも、他の評価視点との共起する確率と比較して特定の評価視点とだけ共起する確率の高い語には大きな関連度を持たせるためである。例えば、評価視点「デザイン」の場合「スタイリッシュ」という、評価視点「デザイン」と共起する確率が低い語でも、評価視点「デザイン」を含む文にしかこの語が登場しなければ「スタイリッシュ」という語は「デザイン」と高い関連を持つ関連語であると考えられる。このような関連語にも高い関連度を持たせるため、対数をとった。

ただし、経験的に出現回数が 1 回の語は関連語として扱わず関連度を持たせないことにする。

3.3 文の言及度算出と分類

3.3.1 言及度算出

すべての文に対して、各評価視点に対する言及度を求める。長文のほうが、具体的に評価が記述されていることが多い。例えば「デザインが良いです。」という意見文と「デザインはシンプルで特に飾り気はありませんが、黒を基調にしているためか安っぽくは見えません。」という意見文があったとする。両者ともデザインのことについて言及しているが、前者より後者のほうが具体的に記述されていて、ユーザは商品購入の参考になりやすい。そこで、文の言及度のスコアリングは後者のほう

(注4): <http://mecab.sourceforge.net/>

が高い言及度を持つようにするために、文に出現する関連語の関連度の総和を言及度とした。

3.3.2 文の分類

3.3.1 節で求めた各評価視点に対する言及度を用いて意見文を分類する。文が持つ各評価視点に対する言及度の中で一番スコアが高い文をその評価視点に対する意見文とする。二番目に高いスコアが負の値をとる場合は、二番目に高いスコアの絶対値が一番高いスコアに加える。これは、一番目のスコア以外が負の言及度を持つ場合、その文が一番高いスコアの評価視点に対する意見文であるという特定性が高いと考えたためである。言及度が負のスコアを持つということは、文に現れる関連語の関連度の総和が負になることである。負の関連度の関連語は、他の評価視点の語と共起する確率が高い語である。関連度の総和を言及度としていることから、言及度が負となる文は、他の評価視点に対して言及しているということが分かる。このことから、一番高いスコア以外のすべてのスコアが負の場合、一番スコアが高い評価視点以外の評価視点に対しては、言及していない文であることが分かる。よって一番高いスコアの言及度の評価視点に言及しているという特定性が高まるため、このような方法を用いた。ただし、二番目のスコアが 0 以上の場合は、二番目のスコアの評価視点に対しても言及しているとシステムが判断しているため、一番目のスコアと二番目のスコアを足さない。

どの評価視点に対しても、言及度を持たないものは分類しない。例えば「昨日購入しました。」という文が、どの評価視点に対しても関連語を持っていない場合、各評価視点に対する言及度は 0 である。このようにすべての評価視点に対して言及度が 0 の場合は分類しない。

4. 実装

本手法を評価するためにレビューサイトとして実際に運用されている価格.com のレビューを使用し、システムを実装した。価格.com のレビューは「評価項目」として、製品に対する主な評価軸が用意されているためである。例えば、価格.com のノート PC の評価項目では「デザイン」、「処理速度」、「グラフィック性能」、「拡張性」、「使いやすさ」、「携帯性」、「バッテリー」、「液晶」が用意されている。このように価格.com のレビューはドメインごとに、ユーザが点数を与えて評価する評価項目が用意されている。この評価項目には、ユーザが与えた点数も付与されていることから、評価項目を評価視点としてレビューを分類することで、点数の根拠となるレビューを得ることができる。価格.com にはユーザが点数を与えて評価できる「満足度」という項目もあるが、「満足度」は対象となる商品全般について述べているもので、商品の側面でないため、本論文では評価視点として利用しない。

提案手法は Java を用いて実装した。入力には価格.com が用意した同じ評価視点を持つドメインのレビューと、価格.com が用意したそのドメインの評価視点とした。同じ評価視点を持つドメインは、例えば「ノート PC」や「液晶 TV」や「洗濯機」などである。同じ評価視点を持つドメインのレビューは、関連

語辞書を作成のためのレビューと、分類の対象となるレビューの 2 つのレビュー集合を用意する。出力は、分類の対象となるレビューからシステムがある評価視点に対する意見文と判断した文をその評価視点に分類した結果である。

5. 評価

4. 章で実装したシステムを使って、提案法の上位 N 位の精度と再現率を求め、従来法の SVM を使った分類と精度を比較する。ドメインはノート PC とし、分類の指標とする評価視点は、4. 章で述べた、ノート PC に用意されている 8 つの評価視点とする。コーパスはノート PC の 2009 年 7 月 17 日での新着レビュー 3000 件を用いる。その中から人手で文がどの評価視点について言及している意見文なのかを判定した 1220 文をテストデータとして評価を行う。

次の 3 パターンのデータセットで辞書を作成した。

A: テストデータ以外のすべてのレビューテキストで作成

B: テストデータも含めた 3000 レビューすべてのレビューテキストで作成

C: 603 文で作成

3 つの辞書を使って、テストデータを分類し評価を行った。

評価に使ったデータを表 1 に示す。以降、テストデータ以外のレビューテキストで作成した辞書を辞書 A、すべてのレビューテキストで作成した辞書を辞書 B、テストデータ以外の正解判定を行った文で作成した辞書を辞書 C とする。

表 1 関連語辞書作成に使ったデータ

| データ | 文の数 |
|----------------------------|---------|
| テストデータ以外のレビューテキスト | 35341 文 |
| すべてのレビューテキスト | 36561 文 |
| テストデータ以外の正解判定を行った文（正解判定なし） | 603 文 |

5.1 精度と再現率

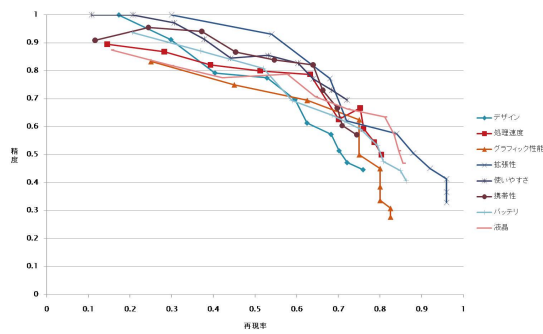
本手法は、文が言及度を持つため、どの程度評価視点について言及しているのかを求めることができる。システムは言及度が高い順番に並べて提示するため、上位 N 位の精度と再現率を求めた。上位 N 位として、分類された文数の 10%, 20%, ..., 100% の数の文を上位から取り出し、精度と再現率を求めた。分類された文がその文について言及している文かどうかの判定は人手で行った。

精度と再現率の計算は以下のように行った。

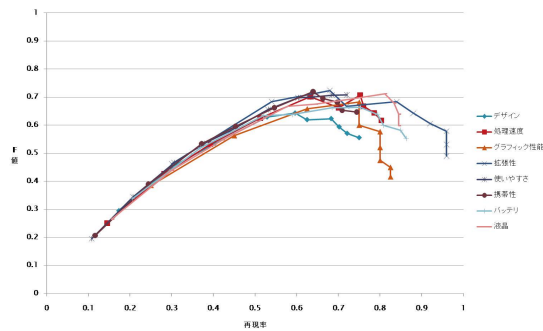
$$P_V = \frac{\text{システムが評価視点 } V \text{ に分類した中で正しく分類された文数}}{\text{システムが評価視点 } V \text{ に分類した文数}} \quad (2)$$

$$R_V = \frac{\text{システムが評価視点 } V \text{ に分類した中で正しく分類された文数}}{\text{評価視点 } V \text{ に分類される正解文の数}} \quad (3)$$

P_V は評価視点 V における分類の精度を、 R_V は評価視点 V における分類の再現率を表している。式 3 の分母は、人手でテストデータを分類した時に、評価視点 V に分類された文の数である。式 2 と式 3 の分子は、システムが分類した文を人手

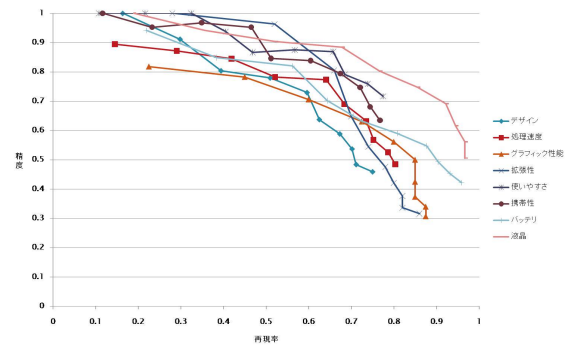


(a) 精度と再現率

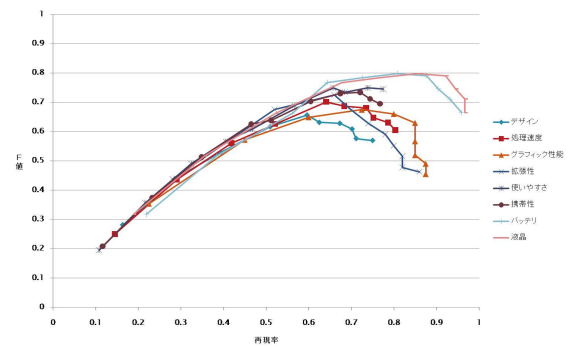


(b) F 値と再現率

図 3 辞書 A を使った分類の結果

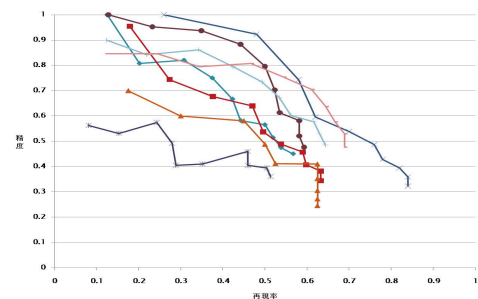


(a) 精度と再現率

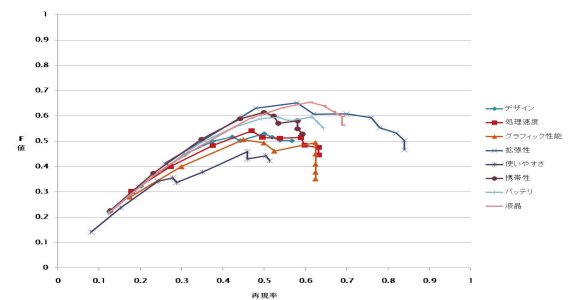


(b) F 値と再現率

図 4 辞書 B を使った分類の結果



(a) 精度と再現率



(b) F 値と再現率

図 5 辞書 C を使った分類の結果

で正解判定し、正しく分類された文の数である。上位 N 位の精度と再現率は、システムが評価視点 V に分類した文数を、上位 N 位までに分類された文の数として求めた。

図 3(a) は辞書 A, 図 4(a) は辞書 B, 図 5(a) は辞書 C を使ってテストデータを分類した精度と再現率を示す。図の横軸は再現率, 縦軸は精度であり, グラフは評価視点ごとの分類結果である。

再現率が 1 に届かない理由として, 意見文が正しい評価視点に分類されず他の評価視点に分類されたからである。例えば, 「デザインが良いです。」という文が「処理速度」の評価視点に分類された場合, 評価視点「デザイン」のところでは再現率が低下する。

図 5(a) は図 3(a) や図 4(a) のグラフと比べて評価視点ごとに精度と再現率が大きく異なっている。これは評価視点の語を含む文の抽出した数が影響している。本手法では辞書を作成する際に評価視点の語を含む文を抽出し, その抽出した文を使って関連語と関連度を求めている。そのため, 抽出する文の数が少ないと, 関連語の数が少なくなり, 関連度は文の集合全体での出現回数や評価視点の語と共起する回数が 1 回違うだけで大きく変わる。図 5(a) の「使いやすさ」の精度と再現率は, 再現率の向上とともに精度も向上している部分や, 精度が極端に低下している部分がある。これは, 大きな関連度を持つ 1 語によって言及度がほぼ決定されてしまい, その語が正解文に出現する確率による影響を受けているためである。

5.2 機械学習による分類と比較

分類の精度評価を行うため、比較対象として機械学習手法の SVM を使って実験を行った。SVM はツールとして Lib-SVM^(注5)を使用し、デフォルトの RBF カーネルを用いて multiclass 分類を行った。特徴ベクトルの作成は本手法と同様の素性を用いた。重みは、バイナリで与えた。他のパラメータはデフォルト設定のまま使用した。データは表 2 を使い、分類の精度を測った。学習データは辞書 C を作成したときに使用した 603 文を、人手で正解判定をしたものとした。レビューの中には、「昨日購入しました。」など意見文でない文や「ファンの音がうるさいです。」など与えた評価視点に対する意見文でない文も含まれる。そのような文は、各評価視点に対する意見文ではないとする「その他」というクラスを用意し、ノート PC の評価視点の 8 クラス+「その他」の 9 クラスで分類した。

本手法では、文が持つ各言及度の中で一番高い評価視点に文を分類したが、5.1 節のグラフから、言及度が低い文は、別の評価視点について言及している文であったり、どの評価視点に対しても言及していない文が含まれていることが多い。図 3(b)、図 4(b)、図 5(b) に縦軸に F 値、横軸に再現率をとったものを示す。F 値が最も高い点の言及度を閾値として、その閾値を超えなかった言及度の文はすべて「その他」の評価視点に、つまりどの評価視点に対しても言及していない文として分類し精度を求めた。

表 2 SVM に使用したデータ

| データ | 文の数 |
|--------|--------|
| テストデータ | 1220 文 |
| 学習データ | 603 文 |

前節で述べた 3 つ文の集合で作成したそれぞれの辞書を使って分類を行った精度と、SVM による分類の精度の結果を表 3 に示す。SVM の精度と比較して、テストデータ以外のレビューテキストで作成した辞書 A を使って分類した場合は、精度が 4.9% 向上し、すべてのレビューテキストで作成した辞書 B を使って分類をした場合は、精度が 6.9% 向上した。学習データに使った 603 文で作成した辞書 C を使って分類した場合は、精度が 1.9% 下がった。

表 3 機械学習法との比較

| 手法 | 精度 (%) |
|-----|--------|
| 本手法 | A 72.7 |
| | B 74.7 |
| | C 65.7 |
| SVM | 67.8 |

表 4 「音」に分類された文

| 文 | 言及度 |
|------------------------------------------------------------------------------------------|-------|
| スピーカーがモノラルとわかっていましたが、最大にしても音楽や映画の音がすごく小さいです | 16.38 |
| あまり音質はこだわらないので、モノラルスピーカーでもよいかと思いましたが、音悪すぎなのにビックリしました | 15.68 |
| ダイヤル式のボリューム調整は便利ですが、無音と最小の間にもう一段階あるとうれしかったです | 13.73 |
| 耳を傾けて 15cm くらいまで近づけばファンの動作音がかるうじて聞こえるレベルです | 12.39 |
| そのため、ファンが断続的に回転して、静かな場所では結構気になるかもしれません | 11.07 |
| 一番心配していたのは静音だったのですが、かなり静かです | 10.40 |
| 前の PC がザーザーと大きな音を立てながら動いていたのに、サーと軽快な音で動いています | 10.38 |
| ファンが殆ど回らないので静か | 9.19 |
| 職場では PHOTOSHOP、メール他事務的作業など、家では iphone をつないで音楽のダウンロードやブログ、ネットショッピング、写真の整理やネットをするのに使用しています | 8.96 |
| 最近、CPU のファンが回り出しましたが、静かなので気になりません | 8.23 |

5.3 評価視点の追加

ノート PC に与えられている 8 つの評価視点に新たな評価視点を加え 9 つの評価視点に基づいて分類をした。加えた評価視点は「音」という評価視点とした。この「音」という評価視点は、価格.com が用意していない評価視点である。この評価視点を加えて分類し、「音」に分類された結果の上位 10 文を表 4 に示す。分類した文は 8 つの評価視点での分類と同様にテストデータの 1220 文である。表 4 から「音」に分類された文はノート PC の音質や、静音性が記述されている文が分類されていることが分かる。

6. 考 察

評価視点別の意見文分類について、評価視点ごとに関連語辞書を作成し、関連語辞書を用いて算出した言及度により意見文を分類した。その結果、本手法による分類精度は約 75% であり、SVM 法と比較して 7% 程度精度が向上できることが分かった。評価に使った辞書は前章で述べた 3 つのパターンで作成したものであったが、603 文で辞書を作成した時は精度が SVM より 2% と低い値となった。このことから、ある程度十分な量のコーパスを学習させ、辞書を作成しなければならないことが分かる。辞書を作成するときに必要となるコーパスの量を検討することは重要である。

評価視点の追加については、ユーザが新たに評価視点を追加しても分類できることが確認できた。追加による精度と再現率への影響に対する評価は行っていないため、今後評価する必要があると考えられる。

前章のグラフから本手法では再現率が 1 にならないことがわ

(注 5): <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

かる．一文が一つの評価視点に分類するだけでなく，文の各評価視点に対する言及度のなかで二番目や三番目の言及度にも着目して，それに応じて一文を複数の評価視点に分類させることもできるようにすることで，精度と再現率を向上させることができると考えられる．

評価視点の語を含まない文も抽出できることを確認できた．例えば「デザイン」の評価視点では「色は落ち着いたある「ブラウン」を選びましたが、「ホワイト」や「ブラック」のように指紋などの汚れが付きやすい光沢のある表面材質を避けたことが最大の理由ですが、大正解でした」が上位2位「色は赤なので派手かなと黒と迷いましたが、ワインレッドなので意外と落ち着いた感じが、シックな感じです」という文が、上位6位に分類されていた．このように評価視点の語を含んでいない文でも，評価視点について詳しく言及している文が高い言及度を持つことができ、正しく分類できた．言及度でみると，上位に分類されたものでは，評価だけではなくその理由が述べられているものが多く，下位のほうでは「デザインがいいです。」など，単に評価を記述した文など詳しく述べられていない文が多かった．このことから評価視点について詳しく言及している文が高い言及度を持つことができ，上位に分類できていることが分かった．しかし，言及度の定量的な評価は行っていないため，今後評価をする必要がある．

関連語辞書では，出現回数がかなり低い語でも高い関連度を持たせることが可能であるが，評価視点と関係ないと思われる語も高い関連度を持つことがあった．これは偶然評価視点の語と共起し，他の文の中では現れない語であった場合にこのように高い関連度を持ってしまつたためである．今回は，出現回数が1回の語は関連度を持たせなかったが，2回や3回の語も出現回数が十分に低いと考えられる可能性があるため，関連度を持たせなくする出現回数の閾値を変化させ，検証することは重要である．

7. ま と め

本論文は，評価視点に基づいて，意見文进行分类することを目的として研究を行った．具体的には，評価視点ごとに関連語辞書を作成し，関連語辞書を用いて算出した言及度を用いて意見文进行分类した．関連語辞書は，評価視点の語と共起する語を関連語とし，関連語が評価視点と共起する確率と，全体での出現確率を使って，関連語ごとに関連度を登録しておく．言及度は，各評価視点の関連語辞書を使って，文に現れる関連語の関連度の総和をその評価視点に対する言及度とした．

実運用されているレビューサイトの評判情報を用いて提案法を実装評価した．従来法のSVMを使った意見文の分類法と比較し，本手法の有効性を確認した．本手法による分類の精度は約75%であり，SVM法と比較して7%程度精度が向上することが分かった．提案法により「ヘアライン加工の天板がいい感じです。」が「デザイン」という評価視点に「やはり700g代は軽いです。」が「携帯性」という評価視点に分類できるなど，評価視点の語を含まない文も正しく分類されていることが確認できた．

今後の課題は，他のドメインでの評価や，価格.comが用意してある評価視点以外の評価視点をユーザが与えた場合のさらなる評価が必要である．他にも，どの程度評価視点に対して詳しく述べている文かどうかの言及度そのものに対する評価も行う必要がある．関連度の計算手法についても今後の課題である．文には「デザインが良く、バッテリーも長持ちです。」のような一つの文で複数の評価視点に言及している文がある．実装した分類手法では，一つの文は一評価視点に分類するとしているが，言及度が最も高い評価視点に分類するだけではなく，2番目や3番目に高い言及度が相対的に高い言及度であればその評価視点にも分類するなど，複数の評価視点に言及している文の扱いについても検討していく必要があると考えられる．

文 献

- [1] 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. pp. 201-241. 2006
- [2] Minqing Hu and Bing Liu. Mining aMinqing Hu and Bing Liu. Mining and summarizing customer reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 168-177, 2004.
- [3] 笈島郁子, 嶋田和孝, 遠藤勉. 系列パターンを利用した評価表現の分類. 言語処理学会第11回年次大会, C2-9. 2005.
- [4] 立石健二, 福島俊一, 小林のぞみ, 高橋哲朗, 藤田篤, 乾健太郎, 松本裕治. Web 文書集合からの意見情報抽出と着眼点に基づく要約生成. 情報処理学会研究報告, 2004-FI-076, Vol. 2004, No. 93, pp. 1-8, 2004.
- [5] 赤木 法生, 大島 裕明, 小山 聡, 田島 敬史, 田中 克己. レビューページ例からの属性抽出に基づくレビューページ検索. 電子情報通信学会第16回データ工学ワークショップ (DEWS2006), 2C-i10. 2006.
- [6] 瀬藤 亮, 佐藤 哲司. 商品説明ページを用いた評価視点別評判情報提示システム. 電子情報通信学会第19回データ工学ワークショップ (DEWS2009), 6C-5. 2009.
- [7] 谷本 融紀, 太田 学. 評価表現辞書を用いた評判情報の可視化. Web とデータベースに関するフォーラム (WebDB Forum2009), 1B-1. 2009.