

# Blogを対象とした商品価値の分析実験

青木 優, 山本 和英

長岡技術科学大学 電気系

E-mail:{aoki,ykaz}@nlp.nagaokaut.ac.jp

## 1 はじめに

近年、インターネットの普及により多くの人が様々な情報を発信することが容易となった。特に Blog や電子掲示板等は書き手の主観的な意見や評価が簡単に入手できることから、マーケティングでは新たな情報源として注目されている。このような情報源から有用な情報を抽出するため、テキストマイニングの分野では様々な手法が提案されている。

Blog を対象としたテキストマイニングではノイズの多さ、辞書の作成、複数の話題への適用等の問題が挙げられる。抽出した意見情報は肯定的/否定的などの二極指標で分類されることが多く、商品の機能や性能に関しての情報がほとんどであるが、商品が提供する価値感について言及されていることは少ない。

そこで我々は消費者の価値観という点から肯定的/否定的の判断に加え、意見情報に含まれる機能的/情緒的価値の分類を行った。商品の価値を表現する指標としては機能や性能を示す機能的価値、自身の感情を示す情緒的価値が考えられる。本稿では Blog を対象とした商品価値の機能的価値、情緒的価値の分類の可能性について検討を行った。

## 2 本研究の位置付け

### 2.1 消費者の価値への注目

丸岡 [1] は消費者の価値観に注目し、ブランドや商品が持つ属性やベネフィットと消費者が持つ価値観の対応を手段目的連鎖モデルによって示している (図 1)。ここで、価値観とは究極の目的であり、情緒的/心理的ベネフィット及び機能的ベネフィットは商品の購買、使用によって得られる情緒面/心理面及び機能面の便益である。属性は商品の持つ特徴を示している。丸岡らの手段目的連鎖モデルでは商品属性から価値観までを抽象化のレベルによって価値観、情緒的/心理的便益、機能的便益、属性の4レベルに分類し、ラダリング調査によって属性 便益 価値観のつながりを明らかにしている。

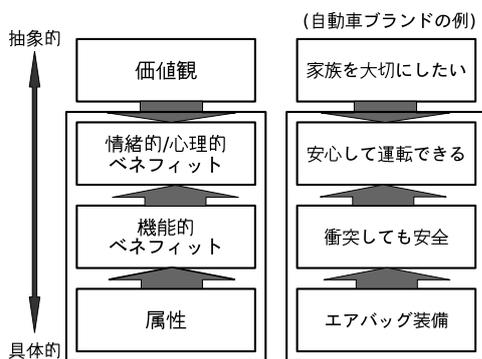


図 1: 手段目的連鎖モデル [1]

既存の手法で抽出された意見情報は機能的便益、情緒的便益が混在しているが、肯定的、否定的の判断しかされていない。なぜこの商品は売れたのか？ なぜ人気があるのか？ 商品の購買及び使用の理由には深層心理である消費者の価値観が背景にある。性能など機能的便益から購入をすることもあれば、感情などの情緒的便益から購入することもある。これらの価値観は異なるもので区別が必要になってくる。

### 2.2 本研究における価値観

価値観は消費者が商品进行评估する判断基準であり、心の中では確かに定められている。人はその基準に従って行動している。そのような心理的要因を直接的に知ろうとするのは困難である。しかし、例えば商品に対する意見や評判情報を分析することから、それを手掛かりとして間接的に消費者の価値観を観測することは可能であると考えられる。

そこで我々は Blog 記事から抽出した意見情報を機能的価値と情緒的価値に分類することで、肯定/否定の分類から一歩踏み込んだ意見情報の分析を行った。本稿では意見情報に対して図 2 に示すような 4 分類を行う。

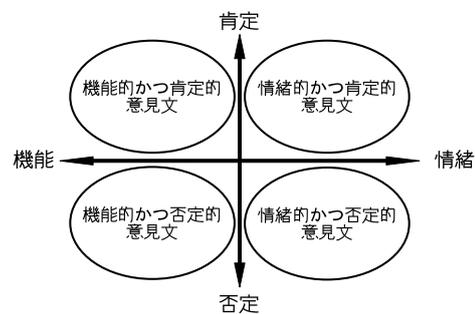


図 2: 意見情報に対する価値分類

本稿における意見情報を以下のように定義する。

- 商品について肯定的もしくは否定的意見が述べられている。(肯定的/否定的意見)  
例 1) 高性能のデジカメに満足満足でした
- 商品の機能や性能、もしくは使い方に関する説明が述べられている。(機能的価値)  
例 2) 夜の室内でもナイトモードなら全然きれい!
- 商品に対する印象や思い出が述べられている。(情緒的価値)  
例 3) 定評のある製品なので満足できると思います。
- 機能的価値と情緒的価値の両方が混在するものに関しては機能的価値とする。  
例 4) この機能でお値段もお手ごろ価格 すっかりデジカメを気に入って来たようでなんとなく僕までうれしくなっていました。

上記に該当しないものはノイズとする。

### 2.3 関連研究

文単位での意見情報抽出としては、属性・評価表現辞書を用いて意見文を収集する研究が行われている。立石ら [2] は意見情報を { 対象物、属性、評価 } の三つ組で表現している。あらかじめ作成した三種類の辞書 (対象物、属性、評価) と抽出ルールを手掛かりに Web 文書から三つ組を抽出し、意見文を判定するルールに従って意見文かどうかの判定を行っている。我々 [3] は評価表現辞書や文書の主題等を手掛かりとして、文単位に意見文スコアを付与している。評価表現は繰り返し学習を行うことで半自動的に収集できる。検索エンジンを用いて文中の主題を判定することで、意見文の判定を行っている。これらの手法はある程度限定

された話題内で意見情報を収集する場合には非常に有効な手法であるが、辞書の構築に手間がかかるという問題がある。

文書単位での意見情報抽出手法としては、意見文を含むコーパスを統計的に処理することで、肯定的/否定的の文書分類を行う研究が行われている。鈴木ら [5] は評価表現や周辺単語を素性としたナイーブベイズ分類器による意見情報分類を行っている。EM アルゴリズムと組み合わせることでモデルを最尤推定しているが、適切なパラメータ推定の問題が残る。藤村ら [4] は肯定的/否定的の文書中の出現頻度を用いることで単語ごとのスコアリングを行っている。各文書に含まれる単語のスコアの総和を求めるとして肯定的/否定的を分類している。しかし、単語単位でスコア付けをしているため意見文以外のノイズまで抽出してしまう可能性がある。また、既存の研究では肯定的/否定的という視点でしか意見情報は分類されていない。機能的、情緒的のような価値観については言及はされていない。

本稿では、あらかじめ分類タグの付与された Blog 記事コーパスを使用し、機械学習による分類手法を用いることで評価表現辞書の構築のような手間を省いた。従来研究では意見情報の収集を行う際、意見文以外のノイズまで抽出してしまう問題がある。これに対し本研究では意見文判定を行いあらかじめノイズを取り除くことでこの問題の解消を図った。これより比較的容易で高精度な分類器の作成を目指す。加えて、肯定的/否定的以外に新たに機能的/情緒的という指標での分類可能性について検討を行った。文書単位の分類では同文書中に肯定的/否定的のような主張の異なる意見文が混在する場合が考えられるため一文単位の分類を行った。

### 3 価値分類手法

本稿での処理の流れを以下に示す。

#### Step 1 分類モデルの作成

機械学習を用いて分類モデルを作成する。

#### Step 2 意見文判定

意見文かどうかの判定を行い、意見文以外のノイズを除去する。

#### Step 3 価値分類

意見文と判定された文を肯定的/否定的、機能的/情緒的で分類する。

#### 3.1 タスク設定

本稿におけるタスクの設定について述べる。本稿ではまず最初に意見文の判定を行い、続いて肯定的、否定的意見または機能的、情緒的価値という指標で意見文と判定された文を分類する。今回、タスクの簡単化のためにこれらを意見文/ノイズ、機能的/情緒的、肯定的/否定的の二値分類問題に分解した。それぞれの問題を解くための二値分類器を用意した。

二値分類器には Support Vector Machine(以降 SVM) を使用した。SVM はパーセプトロン型の教師あり機械学習手法である。マージンの最大化と分離超平面により最適解が唯一に定まり、高精度の分類が可能であるという利点がある。また、藤村ら [4]、鈴木ら [5] の実験で述べられているように、SVM による機械学習は意見情報抽出の問題において彼らの提案手法にも劣らない、高い分類精度が得られている。

#### 3.2 前処理

適切な係り受け解析を行うために Blog 記事を文単位に分割する。前処理として半角の英数字、記号は全て全角に変換した後、以下に示すような文末記号等を手掛かりに文に分割を行った。

- ・ 句点等の文末記号 (。?!?)
- ・ 丸括弧の中身 (笑 泣 爆 涙 汗 怒 苦)
- ・ URL
- ・ リスト形式の記号 ( )

文末記号や文末に存在する丸括弧の中身が以上のような場合その直後で文を改行した。これらは記事を観察した上で経験則によって決定した。記事中には他の Web ページを参照するものも多く見られた。URL が文中に含まれている場合係り受け解析ミスの原因になってしまう。そのため、URL は一文とみなして処理を行った。リスト形式で書かれている記事も見られた。これについてはリストの文頭記号の直前で改行した。

Blog 記事は記事単位で肯定/否定などの分類タグが付与されている。分割された文のタグについては記事のタグをそのまま付与した。

#### 3.3 意見文判定

Blog 記事中には多くのノイズが含まれている。そのため、本稿ではあらかじめ Blog 記事から意見文以外のノイズを取り除くことにした。意見文の判定は意見文であるかどうかの二値分類を行った。意見文を正例、ノイズを負例として SVM に学習させることで意見文/ノイズの二値分類モデルを作成した。本手法では意見文と判定されたものに対して価値分類を行う。そのため、この意見文判定が後の分類精度に影響してくる。

#### 3.4 価値分類

本稿における価値分類とは意見文を肯定的/否定的意見、機能的/情緒的価値の二つの指標で分類することである。肯定/否定の分類については肯定的意見文を正例、否定的意見文を負例として SVM に学習させ、肯定/否定の二値分類モデルを作成した。機能/情緒の分類についても同様に二値分類モデルを作成した。意見文と判定されたものに対しては最終的に、肯定的もしくは否定的、機能的もしくは情緒的の分類タグが付与される。

#### 3.5 素性

学習の際に使用した素性を以下に挙げる。各分類器においてこの 7 種類の素性それぞれについて分類実験を行いその評価を行った。

1. 記号、未知語を除いた形態素全て (1-gram)
2. 記号を除いた形態素の接続 (2-gram)
3. 記号を除いた文節 (segment)
4. 品詞が名詞、形容詞の形態素 (stopword1)
5. 品詞が名詞、形容詞、動詞、副詞の形態素 + 助動詞「ない」 (stopword2)
6. 名詞、形容詞、動詞、副詞、助動詞「ない」の素性の重みを二倍 (weight1)
7. 単語スコアにより重み付けした形態素 (weight2)

1. ~ 5. についての素性の重みはその出現のみを考慮した。記号は意見文としての情報が少ない。未知語は記事の話題によって様々であるためノイズになりやすいと考える。よって、記号と未知語は素性から除いた。

4. ~ 5. については記号、助詞などの情報の少ない語を削除するためにストップワードを設けた。

6. については情報量を削らないまま意見文において重要と思われる品詞の素性の重みのみを上げた。素性には 1. の素性を用いた。

7. については藤村ら [4] の手法を参考に (1) 式より素性のスコアを計算し、素性の重みとした。素性には stopword2 を用いた。

$$score(w_i) = \frac{P_P(w_i) - P_N(w_i)}{P_P(w_i) + P_N(w_i) + k} \quad (1)$$
$$(-1 \leq score(w_i) \leq 1)$$

ここで、 $P_P(w_i)$  は正例で素性  $w_i$  が出現する確率である。同様に  $P_N(w_i)$  は負例で素性  $w_i$  が出現する確率である。また、 $k$  は例えば  $P_P(w_i)$  が 0 であった際、 $P_N(w_i)$  によらずスコアが 1 になってしまう問題を解決するため分母に設けられた実数である。今回は  $k = 0$  とした。

本稿では各分類器において効果的な素性を発見するとともに、どの程度の精度で分類可能であるか調査を行った。

## 4 評価実験

前処理として茶釜<sup>(1)</sup>による形態素解析、南瓜<sup>(2)</sup>による係り受け解析を行った。二値分類器は TinySVM<sup>(3)</sup> を使用し、線形カーネルで実験を行った。SVM のその他のオプションについてはデフォルトのまま使用した。

### 4.1 分類タグ付きコーパス

分類タグ付きコーパスについて説明する。

コーパスの作成はアンケート調査設計、分析経験のあるチェック担当者1名とアンケート自由回答の分類経験のあるアシスタント3名で分類を行った。まずアシスタント2名により全ての文の判定を行う。2名の判定が異なった文を抽出し、残り1名のアシスタントがチェックの上、再判定を行なう。最後にチェック担当者が全ての文をチェックし、誤りがあれば訂正する。

分類タグは { 意見文/ノイズ、機能的/情緒的、肯定的/否定的 } である。分類タグは書き込みごとに付与されている。Blog 記事には宣伝広告や意見性を持たない文などノイズも大量に含まれている。訓練データにはパソコンやデジカメ、携帯電話、ビール、お茶、シャンパーなど、複数の話題が混在している Blog 記事を使用した。これは複数の話題にも対応できるか実験を行うためである。

学習の際、訓練データの正例、負例の文数を揃えることにした。予備実験より、正例、負例の比率を極端に変えてしまうと精度が低下することが確認されたためである。訓練データの内容を以下に示す。括弧内は正例/負例の文数である。

意見文/ノイズ: 16800 文 (8400/8400)  
 機能/情緒: 6800 文 (3400/3400)  
 肯定/否定: 3000 文 (1500/1500)

複数の話題を含む訓練データでどの程度の精度が得られるか実験するため、評価データには訓練データと異なる話題 (菓子) の Blog 記事を使用した。評価データの内容は以下の通りである。括弧内は正例/負例の文数である。各分類器を独立に評価するため評価データは分類器ごとに用意した。

意見文/ノイズ: 1343 文 (563/780)  
 機能/情緒: 461 文 (254/207)  
 肯定/否定: 406 文 (314/86)

### 4.2 各分類器の精度

それぞれの素性を用いた時の各分類器の精度について比較を行った。使用した素性とその精度を表1に示す。

表 1: 各分類器の精度

素性	意見文判定 [%]	肯定/否定 [%]	機能/情緒 [%]
1-gram	67.8	<b>76.6</b>	67.2
2-gram	67.6	75.6	64.4
segment	62.2	59.4	59.4
stopword1	68.0	69.2	67.9
stopword2	69.0	73.6	67.5
weight1	66.2	72.2	<b>69.0</b>
weight2	<b>69.2</b>	<b>76.6</b>	68.1
従来手法	66.4	76.9	66.8

表1の従来手法は藤村ら [4] の分類法を実装した結果である。使用した素性と実験データが異なるため厳密な比較はできないが、本手法の精度を確認する意味で比較を行った。藤村らの分類法については (2) 式に示す。

$$sentence\_score(S) = \sum_{w_i \in S} score(w_i) \quad (2)$$

$$\begin{cases} if \ sentence\_score(S) > 0 \rightarrow positive \\ \phantom{if} \ sentence\_score(S) < 0 \rightarrow negative \end{cases} \quad (3)$$

ここで、 $sentence\_score(S)$  は文  $S$  のスコアで、文中に出現した素性  $w_i$  のスコアの和で表される。素性には名詞、形容詞、未知語を使用した。

表1中の太字は最も精度がよかった素性である。意見文判定には weight2、肯定/否定分類には 1-gram もしくは weight2、機能/情緒分類には weight1 がそれぞれ最も効果的な素性だった。他の分類器に比べ肯定/否定分類は若干精度が良いことがわかる。

### 4.3 システム全体の評価

4.2 節では各分類器を独立に評価した。本節ではシステム全体での評価を行う。肯定的/否定的意見、機能的/情緒的価値の分類は意見文判定で意見文と判定された文のみを対象とした。各分類器には 4.2 節で最も精度の高かった素性を使用した。(意見文判定: weight2、肯定否定分類: 1-gram、機能情緒分類: weight1)

3.3 節で述べたようにシステム全体の分類精度は意見文判定の結果によって変わってくる。システムの精度向上のため意見文判定の際 SVM のスコアに閾値を設定した。SVM スコアの絶対値をとった値が閾値を超えた意見文のみで価値分類を行ったときの精度を測定した。閾値を変化させた時の精度と意見文の抽出数を図3に示す。

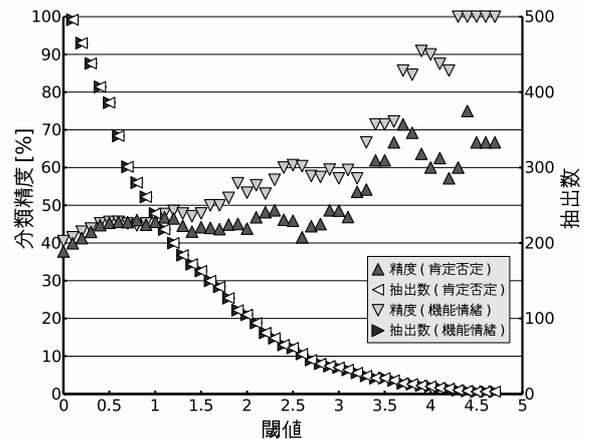


図 3: システム全体の評価

意見文判定の閾値を高くすることで肯定/否定、機能/情緒共に分類精度が向上していることが確認できる。逆に抽出数は閾値に反比例して減少している。

システムが抽出した正解例を以下に示す。

例 5) 分類タグ: 肯定 & 機能

コクのあるミルクキャラメルのおいしさを、そのままチョコレートで再現した親しみやすい味わいです。

例 6) 分類タグ: 否定 & 情緒

.....なんか不味そうだ。

例 7、例 8 に分類を誤った例を示す。

例 7) 誤って情緒に分類された例

ズルズルっ 麺はなかなかシコシコとしておいしいかも。

例 8) 誤って機能に分類された例

チョコレート効果は、以前にも書いたとおり、結構はまっています。

## 5 考察・検討

### 5.1 各分類器の精度について

表1の肯定的/否定的の分類精度を見ると既存の手法に比べて精度が低下している。追加実験として訓練と評価において同様の話題を用いて分類実験を行った。共にデジタルカメラについて

書かれた Blog 記事である。訓練データには 500 文、評価データには 160 文を各分類器ごとに用意した。4.2 節同様に素性を調整した結果、意見文判定で約 78%、肯定/否定の分類で約 86%、機能/情緒の分類で 85%の精度が得られた。その精度は藤村ら [4] の手法に劣らない精度であった。複数の話題が含まれたデータで学習することで異なる話題間でも大幅な精度低下を引き起こすことなく良好な分類結果が得られるが、それはあくまで分類器を独立に評価した場合である。意見文判定と価値分類を組合わせた場合には意見文判定の分類精度が影響し、大きく精度が低下してしまった。決して良好な結果であるとはいえない。

図 3 が示すように意見文判定で閾値を設定することで価値分類の精度は向上する。しかし、閾値を設定することで意見文の抽出数は大きく減少する。それに意見文判定は話題によらず価値分類に比べて精度が低い。システム全体の精度向上にはまず意見文判定の精度を改善することが重要である。これは今後の課題である。

## 5.2 機能的/情緒的価値の分類

機能的/情緒的価値分類について述べる。表 1 より、既存の肯定的/否定的分類手法を用いても機能/情緒分類は良好な精度が得られることがわかった。しかし、肯定/否定の分類と比較すると若干精度が劣っている。我々は意見文を機能的便益が情緒的便益かの二値で分類できると考えて実験を行ってきた。商品の機能や性能等の機能的便益については、従来研究でも述べられているように評価表現辞書だけでは分類が困難である。それに対して評価の主体が人である情緒的便益については、人間の感情を表す表現が出現しやすいと考え、機能的便益に比べて特徴を見出しやすく比較的容易に抽出できると予想していた。しかし、実際には情緒的便益の方が複雑で特徴が掴みにくいものもある。

例 9) 久々のチョコレートは、甘さが歯に染み渡ります

例 9 は「久しぶりに食べたチョコレートがとてもおいしかった。」という感情を表現した情緒的価値の意見文である。この例文には「楽しい」や「うれしい」のような人間の感情を直接表現する特徴的な表現がない。「おいしい」という価値を「歯に染み渡る」という表現に形容しているため、文面の表層情報だけで情緒的価値と判断するのは難しい。

本稿における機能/情緒分類の定義に起因した精度低下もみられた。機能/情緒的価値が両方混在する場合である。

例 10) 印刷もキレイだし、見た目もなかなかスタイリッシュなので気に入っています

例 10 は「印刷がキレイ」「見た目がスタイリッシュ」という機能的価値と「気に入った」という情緒的価値が混在している。本稿では機能的価値と情緒的価値の両方が混在する場合は機能的価値に分類すると定義している。機能的価値を優先しているのは、抽象的な感情表現に比べて具体的な機能や性能の方が商品開発で仕様を変更する際のきっかけとして分かりやすいからである。一方、機械的に処理することを考えたとき、機能/情緒的価値の両方が文中で同時に出現する際は機能とも情緒とも捉えられるような素性が存在し、二値分類器では誤りを起こしやすくなっている。機能/情緒の二値分類では素性が重なることも考えられるためタスクの設定を見直す必要がある。この問題は機能/その他、情緒/その他で独立に分類器を作成することにより、機能的/情緒的価値それぞれの特徴を独立に学習させることで分類精度の向上が図れると考える。これは今後検討すべき点である。

## 5.3 異なる話題への対応

肯定/否定分類ではストップワードを設定し素性の数を減らした場合に精度が低下している。単純に素性が減少したことにより情報量が減ってしまったことが考えられるが、もう一つの原因として訓練データと評価データで出現する名詞、形容詞の傾向が異なることが考えられる。評価表現辞書は限定された範囲の話題で

しか有効ではないように、肯定/否定の表現は話題に大きく依存している。そのため、複数の話題で学習する場合や学習とは異なる話題で評価を行なう場合、品詞を限定してしまうと出現する語の傾向をうまく学習できず精度が低下すると考える。

それに対して機能/情緒分類では素性の数が減っているにもかかわらず若干精度が向上している。これは異なる話題でも情緒的価値に出現する語にある程度の傾向がみられるためである。例 9 のように判定の難しいものもあるが「可愛い」「楽しい」「面白い」など、情緒的価値には出現しやすい語の傾向が見られた。情緒的価値は商品に対する印象や思い入れと定義した。印象や思い入れは人間が感じるものであり、それらの表現にはなんらかの傾向があると考えられる。その傾向をうまく表現できる素性を見つけることができればさらなる精度向上が期待できる。このように機能的/情緒的の分類は素性を工夫することで異なる話題でも対応可能であると考えられる。

## 6 おわりに

Blog を対象した意見情報抽出における分類指標として、機能的、情緒的価値分類の可能性について調査を行った。複数の話題が含まれている訓練データと SVM を用いた機械学習手法により辞書作成の手間を省き、話題に依存しない意見情報の抽出を行った。本稿では、まず意見文判定を行ないノイズとなる意見文以外を除去し、意見文と判定されたものを肯定的/否定的、機能的/情緒的の価値分類を行った。各分類器を独立に評価した結果、素性を調整することで意見文判定は約 69%、肯定/否定の分類は約 77%、機能/情緒の分類は約 69%の精度が得られた。今後の課題として意見文判定精度の向上、機能的/情緒的価値の分類手法の検討が挙げられる。また、機能的価値、情緒的価値をさらに分析していくことで、消費者の価値観と商品の機能や性能との因果関係や、商品に対する消費者の価値観の傾向を調査していく予定である。

### 使用した言語資源及びツール

- (1) 形態素解析器「茶筌」, Ver.2.3.3, 奈良先端科学技術大学院大学 松本研究室,  
<http://chasen.naist.jp/hiki/ChaSen/>
- (2) 係り受け解析器「南瓜」, Ver.0.52, 奈良先端科学技術大学院大学 松本研究室,  
<http://chasen.org/~taku/software/cabocha/>
- (3) SVM 学習ツール “TinySVM”, Ver.0.09, 奈良先端科学技術大学院大学 松本研究室,  
<http://chasen.org/~taku/software/TinySVM/>

### 参考文献

- [1] 丸岡吉人：ラダリング法のブランド戦略への適応，消費者行動研究, Vol.4, No.1, pp.25-39(1996).
- [2] 立石健二, 福島俊一, 小林のぞみ, 高橋哲郎, 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, NL163-1, pp.1-8(2004).
- [3] 峠泰成, 大橋一輝, 山本和英：ドメイン特徴語の自動取得による Web 掲示板からの意見文抽出，言語処理学会第 11 回年次大会発表論文集, pp.672-675(2005).
- [4] 藤村滋, 豊田正史, 喜連川優：電子掲示板からの評価表現および評判情報の抽出，人工知能学会全国大会 (第 18 回)(2004), [http://www-kasm.nii.ac.jp/jsai2004\\_schedule/pdf/000192.pdf](http://www-kasm.nii.ac.jp/jsai2004_schedule/pdf/000192.pdf)
- [5] 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-SWO-A401-02(2004).