

# 意見性判定手法の評価と精度向上

高橋大和 廣嶋伸章 古瀬 蔵 片岡良治

日本電信電話株式会社 NTT サイバーソリューション研究所

{ takahashi.yamato, hirosima.nobuaki, furuse.osamu, kataoka.ryoji }@lab.ntt.co.jp

## 1. はじめに

近年、インターネットユーザへ行動・判断の材料を提供するために、ブログなどの CGM に書かれたクチコミ意見を検索・分析するサービスが登場してきている。

例えば、任意の検索対象に関する評判がどの程度よいかを分析するサービス[1][2]や、あらかじめ限られた対象に関して、評判情報の推移をグラフ化するサービス[3]がある。これらは評判を分析・統計して提示を行うが、本やゲームといった対象に関する好評、不評を表しているブログ記事文を見ることが出来るサービス[4]もある。

このようなサービスを構築するにあたって、重要となるのは、様々な話題を取り上げているブログ記事から、話題に関連した肯定/否定の評判や思想、信念、感情といった書き手の意見を高精度に抽出することである。この基本技術として、われわれは、特徴素性の学習によって意見性を判定する手法[5]を提案し、意見文検索[6]や意見抽出・集計[9]のシステムを実現した。提案の意見性判定手法では、説明文における付加的な主観情報や引用意見を排除して、記事の書き手の意見を文単位で抽出することを目指している。

意見性判定の精度は、効率良く評価や要望、感想といった意見を収集・分析することや、文書分類などを実現する上で重要と考えられる。本稿では、クチコミ意見を検索・分析するサービスの基本技術である意見性判定手法について、さらなる精度向上のための方向性を、特徴素性に関しての評価・分析を通じて考察する。

## 2. 意見性判定手法

意見性判定手法では、ある文において記事の書き手が最も伝えたい部分が主観的であるかどうかで、意見文か非意見文かを判定している。例えば、「会えて嬉しい。」といった、喜びの感情を表している感想や、「この家は美しい」といった評価的意見、また、評判、中立意

見、要望、願望、判断、アドバイス、などを表現した文を意見文として扱う。

意見性判定手法では、以下に挙げる特徴素性を組合せ、意見文の素性ベクトル(正例)、非意見文の素性ベクトル(負例)を SVM により学習することで、正例と負例を分離する分類器を作成する。意見性判定する文に対して、その素性ベクトルを分類器に適用して閾値以上の値が得られれば、意見文と判定する。

### 意見文の特徴素性

- 2,926 意見手がかり表現
- 2,715 意味カテゴリ
- 170 頻出語
- 17 品詞

意見手がかり表現は、意見文を正しく判定するための重要な特徴素性として、収集した意見文から作成されている。意味カテゴリは、日本語彙大系[7]を基にしたシソーラスの階層構造のノードやリーフであり、頻出語は、意見文に出現しやすい特定の単語である。品詞は形容詞や副詞などで、意見性を強める役割があると仮定できるものを、特徴素性としている。

訓練・試験用コーパスは、ブログ検索エンジンを利用して、作品名、人物名、商品名などの検索クエリに関連したブログ記事を集めたものである。ブログ記事本文を文単位に分割して、三人の評価者により、文単位で意見性の有無を付与している。一人でも意見性があると評価した文を正例、三人とも意見性が無いと評価した文を負例としている。クエリ数は 164、記事数は 1,980 記事、文数は 60,257 文である。

このコーパスを、基本訓練用コーパスとオープン試験用コーパス、追加訓練用コーパスに分割した。基本訓練用コーパスとして、72 クエリ分を無作為に選別した。内訳は、正例 8,060 文、負例 15,800 文である。また、オープン試験用コーパスは、321 記事分を無作為に抽出した。内訳は、正例 4,177 文、負例 4,585 文である。追加訓練用コーパスは、残りの 27,635 文で、正例 14,955 文、負例 12,680 文である。

用意した訓練用コーパスから、SVM にて分類器を作成した。以前の報告[8]から、文末から 10 形態素以内に出現する特徴素性のみで判定した方が正解率、正例の適合率が高いことがわかっている。

文末から 10 形態素以内の特徴素性に制限して、基本訓練コーパスのみで学習した分類器をベースラインとして、訓練量による精度への影響を確認する。ベースラインの分類器の精度は表 1 の「基本訓練のみ」の欄に示す。

### 3. 訓練量による精度への影響

分類器は、訓練量を増やすことで、精度が改善されると期待される。ここでは、追加訓練用コーパスを使って、追加訓練を行う。

追加訓練用コーパスを二分割して、追加訓練用コーパス 1 と 2 とした。内訳は、追加訓練用コーパス 1 が正例 6,339 文、負例 7,477 文、追加の訓練用コーパス 2 は、正例 6,341 文、負例 7,478 文である。

表 1 追加訓練と精度

訓練量	正例の適合率	正例の再現率	正解率
基本訓練のみ	78.9%	60.1%	73.3%
基本訓練+追加訓練 1	77.9%	66.4%	75.0%
基本訓練+追加訓練 1+2	78.1%	67.9%	75.6%

表 1 の結果を見ると、正解率に改善が見られるが、正例の適合率はやや下がっていることがわかる。訓練量を増やしたことにより、分類器の作成にかかる時間が増加していることを考慮すると、単純に訓練データを増やすだけでは、効果が薄いことがわかった。

### 4. 誤判定文例による訓練

単純に訓練データを追加するだけでは、意見文と非意見文それぞれを正しく判定できるわけではないことが、前節の実験により示された。そこで、誤判定する文例に着目して、訓練量の増加を抑制しつつ SVM の学習を行うことで、精度向上を試みた。

追加訓練用コーパス 27,635 文に対して、ベースラインの分類器で意見性判定を行った。誤判定している文の内訳は、正しくは正例の文が 4,990 文、正しくは負例の文が 2,064 文であった。これらを追加して、追加訓練を行い、

試験を行った。結果を表 2 に示す。

表 2 誤判定文例の追加

訓練量	正例の適合率	正例の再現率	正例の F 値	正解率
正例と負例を追加	73.8%	70.3%	0.720	73.9%
正例のみ追加	70.7%	78.3%	0.743	74.2%
負例のみ追加	82.8%	48.3%	0.610	70.6%

表 2 から、正例を追加した場合は、主に正例の再現率が上昇するが、適合率は下がる。また、負例を追加した場合は、負例に対して同様の現象が起こる。負例のみの場合、正例の適合率が上がるが、これは、負例を誤判定する割合が減ったためである。

実験の結果から、正例のみを追加すれば、少ない訓練量で、全体の正解率、正例の F 値が上昇し、精度向上が期待できることがわかった。今後は、この分類器を対象に改良を行う。

### 5. 特徴素性の出現位置制限

意見性判定では、文全体の特徴素性を扱うよりも、文末付近の特徴素性のみを扱ったほうが精度が向上する。これは、文末付近に書き手の心情を表す主観的表現が現れるためと考えられる。

そこで、各特徴素性の出現位置制限を変化させて訓練を行い、精度にどのような変化があるかを実験した。各特徴素性のみで訓練を行い、実験を行った結果が図 1 である。形態素数とは、文末から何形態素以内の特徴素性を採用するかを示す。

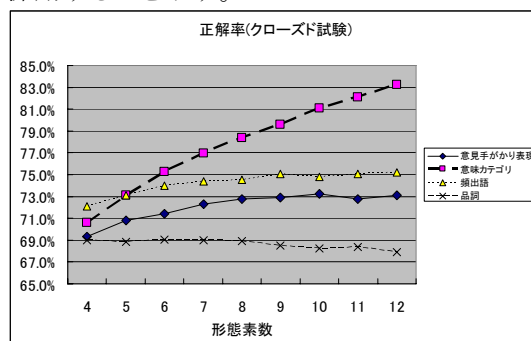


図 1 正解率(クローズド試験)

図 1 から、意見手がかり表現は 9 形態素以上になると正解率の伸びが止まる。頻出語も同様に、9 形態素以上で止まっている。品詞に関しては、6 形態素以上で下がり始めている。

図 2 は、形態素数を変えた場合の正解率の

変化を示している。

意味カテゴリと頻出語は 9 形態素から伸びが止まっている。品詞は一度下がって上がるが、10 形態素からは伸びが止まっている。意見手がかり表現は 12 形態素より増やしても緩やかな伸びが期待できる。

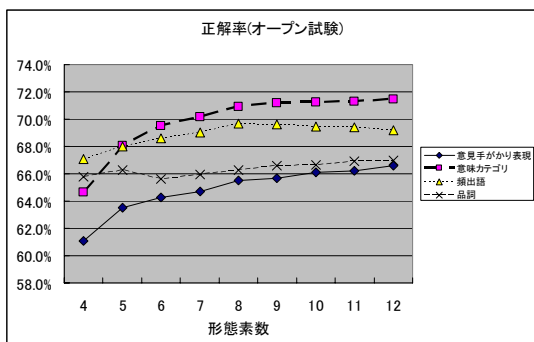


図 2 正解率(オープン試験)

図 3 は、形態素数を変えた場合、正例の適合率がどう変化するかを示している。

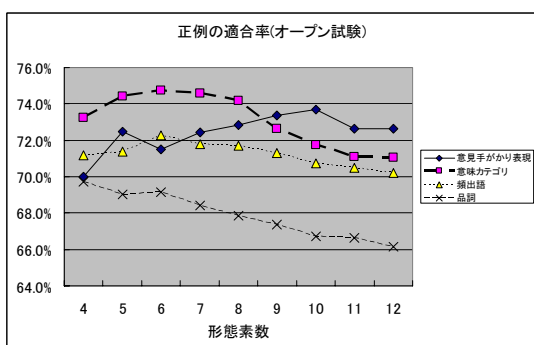


図 3 正例の適合率(オープン試験)

図 3 より、意見手がかり表現に関する結果から、10 形態素以上になると、適合率は下がり始めることがわかる。これは、主に文末に意見性を表す表現が現れるためであると考えられる。しかし、意見手がかり表現は他の素性に比べて正例の適合率への貢献度が高い。また、出現位置に依らない意見手がかり表現もあるため、形態素数は多めにした方がいいと考えられる。

意味カテゴリに関する結果より、正例の適合率に関しては、6 形態素から緩やかに下がっていることがわかる。しかし、正例の適合率は下がるが、正解率は上がることから、意味カテゴリは主に負例(非意見文)に対して効果的と考えられる。

頻出語に関する結果より、正例の適合率に関しては、6 形態素から緩やかに下がっていることがわかる。頻出語は、ある形態素以上になると安定するかやや下がる傾向があることがわかる。

品詞に関する結果より、正例の適合率に関しては、6 形態素から下がり始めていることがわかる。品詞は主に文末付近で効果的な特徴素性であると考えられる。

以上の考察から、頻出語と品詞は文末に近い部分に絞ったほうが良いと考えられる。意味カテゴリは非意見文の判定、また、意見手がかり表現は意見文の判定に効果があると期待できる。よって、出現位置制限として、頻出語は 8 形態素、品詞は 6 形態素以内とし、意見手がかり表現、意味カテゴリに関しては形態素数を変化させて実験を行った。結果を表 3 に示す。

表 3 出現位置制限と精度

出現位置制限	正例の適合率	正例の再現率	正解率
8	72.4%	78.2%	75.4%
9	71.9%	78.6%	75.2%
10	72.2%	78.3%	75.3%
11	71.6%	78.4%	74.9%
12	72.0%	78.5%	75.2%

10形態素の場合で、ベースラインより正解率で 2 ポイントほど向上していることを確認した。

## 6. 負の意見手がかり表現

意見性判定で利用している特徴素性のうち、意見手がかり表現は、特徴的な表現ごとに記述していかなければならない。2,926 種の意見手がかり表現は、基本訓練コーパスから、三人の評価者によって意見文であると評価された文を元に作成している。これらは主に意見文を正しく判定するための特徴素性として作られている。

しかし、分類器の訓練にとっては、意見文、非意見文を判別するための特徴素性として有効であれば、正しい学習が期待できる。非意見文が正しく判別されれば、結果的に意見文の適合率が向上することになる。

そこで、明らかに意見文ではないと判断される文に現れる、非意見性を示す表現を記述し

た意見手がかり表現(以下、負の意見手がかり表現と呼ぶ)を抽出し、それらの表現を特徴素性に追加して、効果を調べる。

該当する文例を収集するため、基本訓練コーパスと、正しくは意見文である追加訓練コーパスに対して意見性判定を行った。結果として、非意見文が意見文と誤判定されている文は、28,850 文中 726 文であった。

誤判定された文例の一部を分析すると、意見を表していると思われる文もあり、評価者による判定エラーと考えられる例もあった。以下に文例を示す。

#### 文例(誤判定された文)

1. これは、もしやすると、大阪は、大変なことになってるかもね。
2. そして、先日、日本語を教えたカナダ人の男の子にお勧め映画を尋ねたところ、『ズーランダー』が面白いよ。とのこと。
3. そんな素敵麻生大臣。
4. よろしかったら～

1.、2.は伝聞(「～になってるかもね」「とのこと。」、3.は皮肉を言っているが、負の手がかり表現では対応しにくい例である。4.も、続く文があれば意見文になりうるため、負の手がかり表現では対応しにくい。

このような文例の一部から、文末付近の表現に注目して、28 種の負の手がかり表現を追加し、再訓練を行った。訓練後、誤判定された文の判定結果は、726 文中 698 文であった。この分類器を使って試験を行った結果を表 4 に示す。

表 4 負の手がかり表現を追加

	負例の誤判定数	正例の誤判定数	正解率
ベースライン	672	1,668	73.3%
正例を追加 出現位置 制限あり	1,259	905	75.3%
正例と負の 手がかり表現 を追加 出現位置 制限あり	1,255	896	75.5%

負の手がかり表現を追加することで、非意見文の誤判定が減少することを確認した。

## 7. おわりに

われわれの提案する意見性判定に関して、誤判定する文例による訓練の追加、4 種の特徴素性の出現位置制限による改良、負の意見手がかり表現の追加を行うことで、一定の改善効果を確認した。

今後は、誤判定する文例を中心に分析を行い、意見手がかり表現と負の意見手がかり表現を整備し、効果を確認する予定である。

また、われわれは、本稿で取り上げた意見性判定手法を応用して、「オピニオン Reader for 映画」[9]を goo ラボにて公開している[10]。本サービスでは、意見文の検索に加え、意見文を好評・不評やイメージワードへ分類し、ランキングを行うことで、映画作品に関する評判の簡便な閲覧を実現している。

#### 参考文献

- [1] blogWatcher  
<http://blogwatcher.pi.titech.ac.jp/>
- [2] BuzzTunes  
<http://www.bztunes.jp/>
- [3] クチコミ評判検索  
<http://buzz.ameba.jp/>
- [4] クチコミサーチ Wipi  
<http://wipi.jp/>
- [5] Nobuaki Hiroshima, Setsuo Yamada, Osamu Furuse, and Ryoji Kataoka. Searching for Sentences Expressing Opinions by using Declaratively Subjective Clues. *Proceedings of COLING-ACL 2006 Workshop "Sentiment and Subjectivity in Text"*, pp.39-46, 2006.
- [6] Osamu Furuse, Nobuaki Hiroshima, Setsuo Yamada, and Ryoji Kataoka. Opinion Sentence Search Engine on Open-domain Blog. *Proceedings of IJCAI07*, pp.2760-2765, 2007.
- [7] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳文, 林良彦. 日本語語彙大系. 岩波書店, 1997.
- [8] 古瀬蔵, 廣嶋伸章, 山田節夫, 片岡良治. ブログ記事からの意見文検索. 情報処理学会研究報告, NL176-18, pp. 121-128, 2006.
- [9] 廣嶋伸章, 古瀬 蔵, 高橋大和, 小田寿則, 富田準二, 片岡良治. 意見抽出・集計システム「オピニオン Reader」. 言語処理学会, NLP2007, 2007.
- [10] オピニオン Reader for 映画  
<http://opinion.labs.goo.ne.jp/>