

述語の選択選好性に着目した名詞評価極性の獲得

東山 昌彦 乾 健太郎 松本 裕治
奈良先端科学技術大学院大学
{masahiko-h,inui,matsu}@is.naist.jp

1 はじめに

1.1 研究背景

近年 blog マイニングなど Web 上における膨大なコーパスの中から意見を抽出し、要約する技術の確立が強く求められている。特にテキスト中に存在する事態が望ましい (以下、ポジティブあるいは p)、望ましくない (以下、ネガティブあるいは n)、またはいずれでもない (以下、ニュートラルあるいは e) の 3 種類の属性値 (以下、評価極性) を自動判定する技術に対する需要が高まっている。

ここで、《素晴らしい/ポジティブ》《酷い/ネガティブ》といった知識が、事態の p/n 判定には有用な情報になりうる。これまでも、Kaji ら [3] や小林ら [12] によって《楽しむ/ポジティブ》《酷い/ネガティブ》といった形容詞または動詞に対する評価極性の獲得は積極的に行われてきた。だが、名詞の中にも以下のように、評価極性を示す表現が存在する。

- あの店の鰹のタタキは 絶品^{ポジティブ} だ。
- 生駒市では近年 不燃ゴミ^{ネガティブ} が増加し、問題^{ネガティブ} になっている。

そこで、このような名詞の評価極性を記述した辞書を構築することを考えた場合、形容詞や動詞に比べ、名詞表現の異なりが膨大なために、人手での作成には多大なコストがかかる。我々はこの問題に対して、名詞の評価極性を自動獲得することで対応することを考えた。本稿では、事態の p/n 判定にとっての名詞の重要性を調べ、p/n に関する述語の選択選好性、教師あり学習、を利用して名詞の評価極性を獲得する手法を提案する。また、提案手法の実験結果を既存研究と比較し、Web コーパス上に頻出する名詞に対しても評価極性の獲得を行なった結果を報告する。

1.2 本論文の構成

2 名詞の評価極性の重要性

我々の目的は、テキストによって伝達される個々の事態情報について評価極性を判別する計算モデルを構築することである。事態を表す言語形式は述語とその項からなる述語項構造と定義する。したがって、事態表現の評価極性を判別する問題にアプローチするには、まず述語項構造の評価極性を判別するための知識・モデルを開発することが先決であると考えられる。述語項構造の評価極性は、述語の評価極性、項の評価極性、述語と項の組み合わせの評価極性など、様々な要因から決まる。そこで、まずは試みに、「名詞+格助詞+用言」の 3 つ組の評価極性がどのように決まるかを調べるため予備調査を行った。

河原らの 5 億文コーパス [6] から「名詞+格助詞+用言」の 3 つ組を網羅的に抽出し、異なりの頻度から推定

される出現確率の分布にしたがい、2,000 件の異なり事例をサンプリングした。これらの事例の評価極性を人手で判定し、さらに述語項構造のどの構成要素が評価極性を決めるかを分析したところ、次のような結果を得た。

- (a) 用言の持つ評価極性で全体の評価極性が決まると考えられる事例 (116 事例, 37%)
例) 試験に合格する^{ポジティブ}, おもちゃが壊れる^{ネガティブ}
- (b) 名詞の持つ評価極性と拡大/縮小を表す用言の組み合わせで決まると考えられる事例 (143 事例, 45%)
例) 説得力^{ポジティブ}が増す^{拡大} (ポジティブ+拡大 ポジティブ), 余裕^{ポジティブ}がない^{縮小} (ネガティブ+縮小 ネガティブ)
- (c) 名詞にも述語にも単体では評価極性が認められないが、組み合わせによって評価極性が生じる事例 (34 事例, 11%)
例) 胃が軽い^{ポジティブ}, 前置きが長い^{ネガティブ}
- (d) 慣用句と考えられる事例 (22 事例, 7%)
例) 舌鼓を打つ^{ポジティブ}, 地に落ちる^{ネガティブ}

ここでの拡大、縮小を表す 増す や ない は単体では評価極性を持たず plus と minus の二つの属性を持つ pn 演算子として乾ら [11] が提案しているものと同一である。pn 演算子は程度の拡大 (plus) や縮小 (minus)、開始 (plus) や中止 (minus) を表す。

確定的な結論を出すにはより大規模な調査が必要であるが、この予備調査の結果は次の点で興味深い。まず、タイプ (b) は名詞の評価極性と pn 演算子から評価極性が構成的に決まるものであるため、タイプ (a) と合わせると、全体の 8 割以上の事例が個別の語の評価極性から構成的に決まることが分かる。一方、タイプ (d) の知識は辞書的知識としてある程度人手で書き尽くせると考えられるので、名詞と用言の組み合わせが問題になるのは実質的にはタイプ (c) の 11% だけである。さらに、(a) の用言にはサ変名詞のように名詞として機能する語や「純朴な・純朴さ」のように名詞形が存在する語が少なくないことを勘案すると、まずは名詞の評価極性に関する知識を一定の規模で蓄積することが重要であることが示唆される。

3 名詞の評価極性の定義

名詞の評価極性の獲得について述べる前に、名詞の評価極性の定義を行う。まず、用言は事態であるから、それらの評価極性はその事態が望ましいか否かを表す評価によって決定される。一方、名詞については以下の3種類に準じる述語を補うことで事態化し、その事態に対して望ましいか否かを判別することで、評価極性を判定する。

1. (~する)を付与して動詞化
2. (~だ)を付与して形容詞化
3. (~が増える)や(~がある)を付与して動詞化

1. についてはサ変名詞に対するものであり、評価極性は動詞に対して適用した場合と同様に決定される。2. については名詞が形容動詞語幹の場合である。これも評価極性は形容詞に対して適用した場合と同様に決定される。3. の場合は2節で述べた、pn 演算子を付与して評価極性を決定する。例えば「信頼」のように がある とおいう plus 演算子を付与してポジティブであれば、ポジティブな名詞であり、「ミス」のように が減る という minus 演算子を付与してポジティブであればネガティブな名詞であると判別する。

4 提案手法

名詞の評価極性知識の獲得のためにまず以下に示す基本的なアイデアに注目した。要素の評価極性についてセンシティブな選択選好を持つものが相当数存在する。例えば、「心がける」のヲ格には「安全運転」や「心配り」のようなポジティブな行為を表す名詞が出現しやすいし、「防ぐ」のヲ格には「火事」、「事故」のようなネガティブな出来事を表す名詞が出現しやすい。この性質を利用すれば、名詞と述語の共起行列から名詞の評価極性を獲得することができる。評価極性の識別に有用な述語は、ある程度の規模の訓練事例があれば、教師あり学習で収集することができる。これを具体化すると概ね次のような手順になる。

1. 共起行列の生成 コーパスから「名詞+格助詞+述語」の3つ組を網羅的に抽出する。この場合、述語については、主辞だけを取るのではなく、「奪われる」のようにモダリティを含めた形で共起を取る。ただし、(a) ガ格、ヲ格、デ格の共起に限定する、(b) 述語とその直前格の共起のみを対象とする、(c) 名詞に修飾語が係っている事例は除外する、などのヒューリスティックな制約を課した。(c) の制約は、「丁寧な言葉遣いを心がける」のように、名詞の修飾語が格要素全体の評価極性を支配する事例を排除するためである。
2. 共起ベクトルの生成 1 で得られる共起行列から各名詞の共起ベクトルを生成する処理にも何通りかの方法が考えられる。後述の実験では、各名詞 N について、それと共起する「格助詞+述語」の組(述語パターン)のうち、(d) N との自己相互情報量が0より大きく、さらに(e) 閾値(5万)以下の頻度となる一般的すぎない述語パターンのみを選択し、対応する素性値を1、それ以外を0とした。このとき、「を防げない」のように否定がつくものについては別の素性として扱う。また、コーパス中での出現頻度が50以下の名詞も評価実験からは除外した。

3. 分類器学習 学習を行うにあたって十分な数のポジティブな名詞およびネガティブな名詞を訓練事例として与え、教師あり学習によって二値分類器を訓練する。実験には、線形カーネルのSVMを用いた。

語の評価極性の獲得については7節に示すようにすでに様々な試みが報告されているが、一般の名詞を対象としたものはそれほど多くない。その中では、辞書の語釈文を利用する Takamura らの手法 [9] が注目される。Takamura らは、辞書の語釈文やシソーラス、コーパスの情報を利用して、異なる語の間の評価極性の一致・不一致に関する局所的依存関係を獲得し、統計的な最適化によって単語集合全体の値割当てを行う方法を提案している。

この手法に対し、述語の選択選好を利用する方法に期待される利点はいくつかある。第1に、述語と名詞の共起を利用することで、Turney ら [10] のような windowing に頼る方法に比べてより精密な文脈ベクトルを与えることができる。また、それと同時に、主節-従属節の共起などに比べるとより大規模に共起データを収集することができる。第2に、述語パターンを素性とする共起ベクトルが名詞間の意味的類似度(すなわち分布類似度)を推定する良い手がかりになることはすでによく知られている [13]。我々の方法はこうした手がかりから得られる名詞間の意味的類似度の情報を評価極性分類に自然に利用することができ、語釈文やシソーラスのような固定的な情報に頼る必要がない。

5 実験

5.1 訓練・評価用データ

手法の有効性を評価するための予備的な実験として、既存手法を代表する Takamura らの手法 [9] との比較を行った。具体的には、Takamura らが公開している単語感情極性対応表のうち、確信度の高い一部の名詞を訓練事例として利用し、4節で示した手法で訓練した評価極性分類器を使って残りの名詞を再分類した。

まず訓練用データとして、Takamura らの単語感情極性対応表からスコアが+0.99以上の名詞をポジティブな訓練事例として、-0.99以下の名詞をネガティブな訓練事例として抜き出した。また、我々のグループで別途作成した評価値表現辞書 [8] に含まれる形容詞、形容動詞の名詞形をこれに加えて、人手でクリーニングし、最終的にポジティブな訓練事例 421 件、ネガティブな事例 692 件を得た。また、Takamura らの辞書の中からランダムにサンプリングし、人手でチェックを行い 848 件のニュートラルな訓練事例を追加した。

評価用データについては、単語感情極性対応表の残りの名詞をコーパス中の出現頻度に従ってランダムにサンプリングし、人手で分類することにより、ポジティブな事例 683 件、ネガティブな事例 441 件、ニュートラルな事例 3614 件を得た。

5.2 実験方法・評価

本稿で提案した手法の有効性を調べるために、Takamura らの手法をベースラインとし、我々のモデルと比較する。我々の述語選択選好性に基づくモデルは、SVM を利用しており、2値分類しか行うことはできないが、ポジティブ、ネガティブ、ニュートラルの3値に対して、ポジ

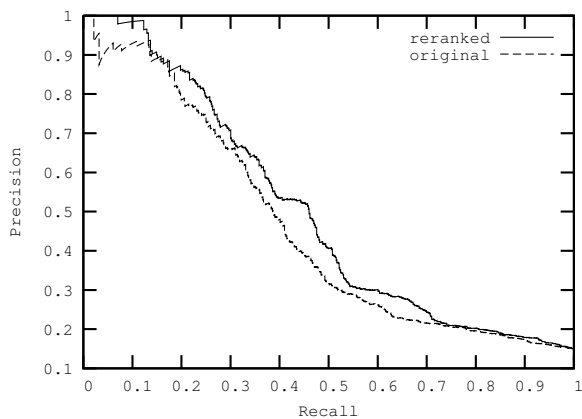


図 1: ポジティブな名詞獲得の P-R 曲線

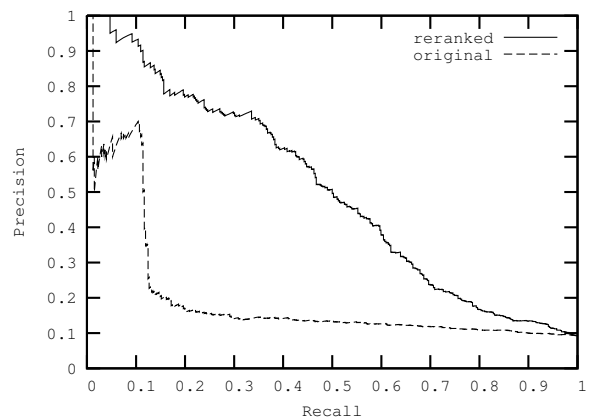


図 2: ネガティブな名詞獲得の P-R 曲線

タイプとそれ以外, ネガティブとそれ以外に訓練事例を分割し, それぞれについて学習・分類を行う. 対象名詞集合を超平面からの距離によってランキングし, P-R 曲線と F 値を用いて比較を行い評価する.

5.3 実験結果・考察

図 1, 図 2 はそれぞれポジティブとネガティブの Takamura らによるランキングと我々のランキングから得られる P-R 曲線を描いたものである. ポジティブ, ネガティブのそれぞれの曲線について, 我々の提案手法が上まわっていることが読みとれる.

図 2 のネガティブの場合は, 極端な差が生じているが, これは Takamura らの手法では, 「猿」「目」「馬」といったニュートラルの名詞が強くネガティブと判定されているためである. それに対して我々の提案手法は比較的ノイズが少ない状態で多くのネガティブな名詞を獲得できている.

F 値が最大となる点による比較を表 1, 表 2 に示した. baseline は Takamura らのスコアを用いた結果であり, SVM は我々の提案モデルを用いた結果である. SVM(一直前格) は, 我々のモデルから共起の抽出関係を直前格のみに限るという条件 [4 節 1(b)] を解除した場合であり, SVM(-修飾) は名詞が修飾されている共起関係のフィルタリング [4 節 1(c)] を行わないものである. 下 2 段はそれぞれ, 述語パターンのフィルタリング条件についてのものであり, SVM(-相互情報量) は名詞と述語パターンの相互情報量が 0 以上という条件 [4 節 2(d)] を除いたもの, SVM(-頻度) は述語パターンに対して出現頻度によるフィルタリング [4 節 2(e)] を行なわなかったケースである. この結果から, 一般的な名詞を足切りすることが強く影響を与えていることがわかる.

また, 表 4 に素性として用いられ, 学習によって強く重みが付与された述語パターンを示す. ここで注目したいのは「を克服する」や「を治す」といった極性反転子 [7] が有効に活用されている点である.

6 Web 頻出名詞の評価極性知識獲得

以上の実験では, Takamura らの辞書との精度比較を行うために判別対象の語を国語辞典中の名詞に限っていた.

しかし, 平均的な国語辞典程度の規模であれば, すべての単語について評価極性を人手で付与することも非現実的では必ずしもない. 評価極性獲得の真の価値は, Web のようなオープンな文書集合に出現するオープンな語彙に対して適用できるかどうかで決まることを考えるべきであろう. そこで, 我々の提案したモデルが, Web 上の名詞に対しても適用可能であることを示すために河原らの 5 億文コーパス [6] から頻度が高い名詞を抽出し, 評価極性を割り当てた.

表 3 に実際の例として獲得されたポジティブな名詞とネガティブな名詞を示す. こうした例が示唆するように, オープンな語彙に対しても提案手法はある程度うまく働いているように見える. 今後は定量的な評価を行うとともに, 本手法のモデルを評価極性辞書開発の支援ツールとして実際に利用するところまで進めたい.

7 関連研究

評価極性を自動判定する技術についてはこれまでさまざまな研究が行なわれてきた. 特に本研究のような評価表現辞書の構築については語彙ネットワークを利用した手法 [9, 2, 4, 1], 文脈一貫性を利用した手法 [10, 3, 5, 11] などがある.

例えば, Takamura ら [9] の研究では, 少量のシードだけを利用した教師なし学習を用い高精度で辞書の構築を行う手法を示した. だが, Esuli ら [1] の研究でもわかるように, 事前により多くの評価極性の知識があれば, 精度が向上するのは必然である. また, 固有表現のクラス同定などのタスクに比べ, 評価極性の同定はクラス数がたかだか 3 つであり, 今回の実験で用いたように評価表現の同定に必要な数の訓練事例を作成することは, 十分に現実的なことと言える. よって我々は研究方針として, 教師あり学習による評価極性知識の獲得を選択した.

さらに, 語彙ネットワークを利用した手法では, 国語辞典などの既知の言語資源を利用するため新語への対応が難しいが, 本研究では Web コーパスを資源として用いるため, 辞典中のエントリだけに留まらない. Kaji ら [3] の研究では, 評価文が出現しやすい文脈に限定したコーパスを構築することで, 文脈一貫性を利用した Turney ら [10] の手法をさらに洗練させた. だが, 形容詞を中心とした評

表 1: ポジティブな名詞獲得の結果

	Recall	Precision	F-measure
baseline	0.37	0.51	0.43
SVM	0.45	0.51	0.48
SVM(- 直前格)	0.42	0.48	0.45
SVM(- 修飾)	0.40	0.55	0.46
SVM(- 相互情報量)	0.40	0.58	0.47
SVM(- 頻度)	0.50	0.35	0.41

表 2: ネガティブな名詞獲得の結果

	Recall	Precision	F-measure
baseline	0.56	0.13	0.21
SVM	0.46	0.57	0.51
SVM(- 直前格)	0.45	0.50	0.47
SVM(- 修飾)	0.52	0.50	0.51
SVM(- 相互情報量)	0.47	0.50	0.49
SVM(- 頻度)	0.36	0.31	0.33

表 3: 学習によって強く重みが付与された述語パターン

positive	を輩出する で知られる を得られる があふれる 輩出する で知られる で得られる で充実する を誇示する があふれる で使いやすい で関心する を競う で落ち着く で楽しめる をかみしめる を心がける
negative	を克服する で困る が襲う を治す を忘れさせる をくろう で倒れる を嘆く を助長する を和らげる を脱す を痛む がつきまとう をやらかす で怖い を働く を抜く が続出する で苦しむ が嫌いだ を免れる

表 4: Web 上から獲得された評価極性を持つ名詞

positive	スピード感 モチベーション チームワーク 達成感 ポップ チャレンジ精神 リゾート気分 アットホーム 自主性 積極性 景気回復 向上心 開放感 フレンドリー 優遇措置 レベルアップ ピンゴ ヒット曲
negative	歪み 生活習慣病 ニキビ リストラ 文字化け エイズ リバウンド ストーカー 不正アクセス スпамメール ガン いじめ 肌荒れ シワ クーデター 夏ばて 生ゴミ メモリ不足 DV オイルショック ドタキャン

価表現にのみ対象を絞っているため、できごとや経験を含めた名詞を対象とする我々の研究とは一部しか重ならない。

8 おわりに

本稿では述語の選択選好性に着目し、教師あり学習によって名詞の評価極性の獲得を行なう手法について提案した。実験結果からは、述語の選択選好性という観点に基づく仮説は名詞の評価極性に効果的に働くことを確認できた。また、Web 上に頻出する名詞に対して評価極性知識を獲得し、国語辞典中のエントリに留まらない単語に対しても獲得が可能であることを確認した。ネガティブな名詞に関しては劇的な精度向上がみられたが、ポジティブな名詞に対しては Takamura らのものと大きく差はなく、更なる改善が必要だろう。今後の予定としては、よりニュートラルな訓練事例を活用するモデルに改善することによりポジティブな名詞に対して精度を向上させ、獲得された名詞を用いて、述語とその項からなる事象に対して構成的に評価極性の判定を行うことができるようなモデルを構築し、適用できる範囲を拡大していきたい。

9 謝辞

本研究は、文科省科研費特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」の公募研究「経験マイニング: Web 文書からの個人の経験の抽出と分類」(19024057, 代表: 乾健太郎) およびニフティ株式会社から支援を受けた。

また、「Web 上の 5 億文の日本語テキスト」の使用許可を下された情報通信研究機構の河原大輔氏と京都大学大学院の黒橋禎夫氏、「単語感情極性対応表」を使わせて頂いた東京工業大学奥村研究室高村大也氏に深く感謝いたします。

参考文献

- [1] Andrea Esuli and Fabrizio Sebastiani. Pageranking wordnet synsets: An application to opinion mining. In *Proceedings of ACL-07, the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 424–431, 2007.
- [2] Mingqing Hu and Bing Liu. Mining opinion features in customer reviews. In *Proceedings of the 19th National Conference on Artificial Intelligence*, pp. 1075–1083, 2007.

- [3] Nobuhiro Kaji and Masaru Kitsuregawa. Building lexicon for sentiment analysis from massive collection of html documents. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1075–1083, 2007.
- [4] Jaap Kamps, Maarten Marx, Robert J. Mokken, and Maarten de Rijke. Using wordnet to measure semantic orientation of adjectives. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pp. 1115–1118, 2004.
- [5] Hiroshi Kanayama and Tetsuya Nasukawa. Fully automatic lexicon expansion for domain-oriented sentiment analysis. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 355–363, 2006.
- [6] Daisuke Kawahara and Sadao Kurohashi. A fully-lexicalized probabilistic model for japanese syntactic and case structure analysis. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 176–183, 2006.
- [7] Alistair Kennedy and Diana Inkpen. Sentiment classification of movie and product reviews using contextual valence shifters. *Computational Intelligence*, Vol. 22, No. 2, pp. 110–125, 2006.
- [8] Nozomi Kobayashi, Kentaro Inui, Yuji Matsumoto, Kenji Tateishi, and Toshikazu Fukushima. Collecting evaluative expressions for opinion extraction. In *Proceedings of the 1st International Joint Conference on Natural Language Processing*, pp. 584–589, 2004.
- [9] Hiroya Takamura, Takashi Inui, and Manabu Okumura. Extracting semantic orientation of words using spin model. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, pp. 133–140, 2005.
- [10] Peter D. Turney. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 417–424, 2002.
- [11] 乾孝司, 乾健太郎, 松本裕治. 出来事の望ましさを目的とした語彙知識獲得. 言語処理学会第 10 回年次大会, pp. 91–94, 2004.
- [12] 小林のぞみ, 乾孝司, 乾健太郎. 語釈文を利用した「p/n 辞書」の作成. 人工知能学会 言語・音声理解と対話研究会 SLUD-33, pp. 45–50, 2001.
- [13] 藤田篤, 乾健太郎, 松本裕治. 自動生成された言い換え文における不適格な動詞格構造の検出. 情報処理学会論文誌, Vol. 45, No. 4, pp. 1176–1187, 2004.