

文のカテゴリと極性の度合いの推定を行う 評判分析システムの研究

Research on reputation analysis systems
that estimate sentence categories and polarity degrees

大竹 翔馬

Shoma Otake

三輪 誠

Makoto Miwa

佐々木 裕

Yutaka Sasaki

豊田工業大学

Toyota Technological Institute

{sd13404, makoto-miwa, yutaka.sasaki}@toyota-ti.ac.jp

1. 背景と目的

近年、インターネットの普及により社会に情報があふれ、人と物の交流が高速化、広範囲化したことで、消費者の製品に対する目は厳しくなり、本当に消費者の要望に沿った製品でなければ市場で成功することが難しくなっている。このように消費者の動向を分析するマーケティングが重要化したこと、さらにインターネット社会におけるデータ量の増大したことにより自然言語処理を用いた評判分析の重要性が年々増してきている。しかし現在の評判分析システムとして、文章分類をするものや関係のある文章のピックアップをするもの、また極性判定において肯定否定を判定するものが多い。本研究の目的は文毎にどのくらい肯定的か、どのくらい否定的かという極性の度合とそのカテゴリを自動的に判定することである。さらにそれにより、レビュー中の各文が何に関係する文で、どの程度の評価なのかを可視化し、レビューの全体像を直感的に把握しやすくすることが可能な評判分析システムの構築も目的とする。

2. 関連研究

レビューの評判分析という研究は広く行われている[1]。評判分析研究の初期のものに Turney[2]による商品レビューの極性判

定の研究がある。この研究ではレビュー中の副詞、形容詞を含む語句の、“excellent”と“poor”の共起頻度を図ることでレビューが肯定的か否定的かを2値で判定している。また Pang[3]らによる映画レビューを対象とした研究では、3種類の機械学習手法と8種類の特徴を組み合わせて、計24パターン中どの組み合わせが最も極性判定に有効かの調査が行われている。この結果、機械学習の手法としてSVMを用いたものの精度が高い傾向にあることが分かっている。また小堀[4]により楽天トラベルのユーザーレビューの極性を判定し、それをもとにユーザーにサービスなどを推薦するレコメンデーションシステムの開発が行われている。この研究ではユーザーレビューの極性判定のためにSVMを用いており、文の特徴ベクトル化をするための方法としていくつかの素性値のとり方を提案している。

また、極性判定についての評判分析だけでなく、レビューなどの評判情報をカテゴリごとに分類する研究もおこなわれており、Nguyen[5]らによる分類実験では、「食べる」や「おいしい」などの特徴語からその後の対象となる名詞句のカテゴリ分類が行われている。

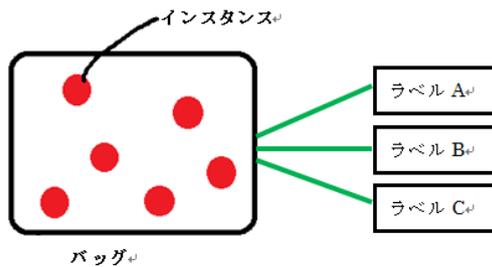


図 1. MIML 学習のイメージ

また近年, Multi Instance Multi Labeling (MIML) 学習 [6] が提案されている. これはマルチインスタンス学習とマルチラベル学習を組み合わせたものとなっている. マルチインスタンス学習は複数の事例が集まってできたバッグに対し一つのラベルが付与されたものを対象とした学習手法であり, マルチラベル学習は一つの事例に対し複数のラベルが付与されたものを対象とした学習手法である. MIML 学習ではこの方法を対象とした学習手法であり, 図 5 のように複数の事例の集まってできたバッグに対し複数のラベルが付与されるというものを対象としている.

3. 提案手法

これまでの研究では評判分析のシステムを用いて, レビューが肯定的か否定的か分類し, その極性を判断することはできたが, その各々の文が, どれくらいの強弱を持った表現で, どのカテゴリについて書かれているかは, 人が読まなければ判断できなかった. そこで, Support Vector Machine (SVM) による文の肯定的・否定的という情報から, 文がどれくらいの強弱を持った表現で記述されているのか, つまり文の極性の度合いを推定する. また, MIML 学習により文ごとのカテゴリを隠れ状態として推定する. さらに, レビュー中の各文が, 何に関係し, どの程度の表現の強さをもった文を使って書かれているのかを可視化する.

4. 実験

本研究では実験データとして楽天トラベルのホテルレビューを用いて提案手法の評価・可視化システムの実現のため以下の 2 つの実験を行った.

1. SVM による文の極性度合の推定
2. MIML 学習によるカテゴリ分類

文の極性度合の推定では 1,200 文に対し人手で肯定的か否定的かを分類し, それを訓練データとし, 学習に使っていない 300 文をテストデータとした. また, カテゴリ分類では各カテゴリの評価値が 5 段階で付与された 100,000 レビューを訓練データとして用い, 500 文をテストデータとして, そのカテゴリの推定を行った.

4.1. 文の極性度合の推定

1) 特徴抽出

訓練データ・テストデータに含まれるレビューを構成する文から特徴抽出を行いベクトル化した. この特徴抽出を行う方法として, 肯定的・否定的な文に各々の特徴的な表現(語彙、言い方)があるとの考え方にに基づき, レビューを構成する語が肯定的・否定的な文どちらに出現しやすいかという情報を特徴とする方法を用いた. 具体的には, 先行研究 [3] で提案された特徴抽出手法の中でレビュー一文に対し SVM による二値分類を行った際に F 値が最も高かったと報告された式(1)を素性値として採用した.

$$\text{素性値 } V(w) = \left(\frac{P(w) - N(w)}{P(w) + N(w)} \right)^2 \quad (1)$$

ここで, w は単語, $P(w)$, $N(w)$ はそれぞれ肯定的・否定的な文中の w の頻度である.

2) 分離平面の作成

上述の方法で特徴を抽出した訓練データを用いて, SVM により分類用のモデル・分離平面を作成した.

3) 強弱の判定

分離平面からのテストデータの各文への距離を算出し、その距離をスコアとして、そのスコアが表現の強弱に相関があるか評価した。ここでいう距離とは SVM における分離平面からの距離である。

評価の仕方としては最も否定的な表現、つまり大きく負のスコアをとったデータから順に表 1 のように 1 から 4 の 4 段階の評価を与えた。そして各文を手で読んで主観により 1 から 4 の点数をつけたものと比較を行うことで一致率を算出した。算出した値は 52% となった。ランダムに評価した場合のベースライン 25% に対して 27% の向上が見られた。表 2 はレビュー文のスコア例である。

表 1. スコアによる評価の付け方

度合	評価	SVM スコアの区間
強く否定的	1	~ -2
否定的	2	-2 ~ 0
肯定的	3	0 ~ 2
強く肯定的	4	2 ~

表 2. レビュー文の SVM によるスコア例

レビュー文例	スコア
コストパフォーマンスが良い	0.47
良いホテルでした	0.889
素敵なホテルでした	0.895
素晴らしいホテルでした	0.913
駅に近くお気に入り	1.00
駅から近く最高	1.34
コストパフォーマンス抜群です	1.95

4.2. カテゴリ分類

図 2 のように、文という事例が複数集まりレビューという集合を作っている状態から MIML により学習し、各文のカテゴリを隠れ状態として推定した。レビューを構成する文が事例、事例が集まってできたバッグがレ

ビュー、レビューの持つカテゴリの評価値がラベルとなっている。文は「立地」、「部屋」、「食事」、「風呂」、「サービス」、「設備・アメニティ」、「その他」という 7 つのカテゴリの全てのカテゴリを持つと仮定し、レビューを構成する各文が持つ各カテゴリの評価値の確率の中で最も高いものをその文のカテゴリとしている。これは最も確率が高いものがこの文の内容の中で最も多くの割合を占めていると考えているためである。実際の文は複数の事柄についての内容を含んでいるが、今回は簡単のため文のカテゴリを一つ選ぶこととした。

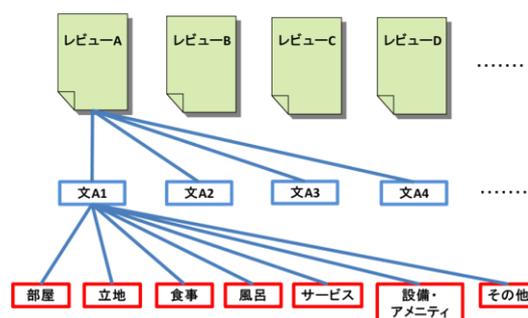


図 2. MIML 学習のイメージ

実験の評価の方法としては、人手で文を読み、各文がどのカテゴリに属するかの正解情報を付与した評価用データと比較し、一致率を計算した。その結果一致率は 35% となり、ランダムに決定した場合のベースライン 14% に対し 21% の向上を得た。カテゴリごとの一致率の詳細を図 3 に示す。

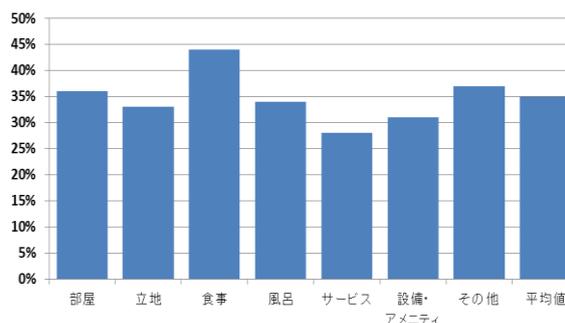


図 3. カテゴリ別の推定の一貫率

5. デモンストレーション

これまでの、極性の度合いの推定と文のカテゴリへの分類を組み合わせることで、文章全体がどのような事柄について書かれていて、さらにどれくらいの強さを持った表現を使って書かれているのかを可視化できるようにした。これによりマーケティングなどの評判分析を行う者がレビューの全体像を一目で把握できるようにするという狙いがある。可視化の仕方としては、各ホテルのレビューをカテゴリごとに色分けし、その色の濃淡で肯定度合・否定度合の強弱を示している。



図 4. 可視化したレビュー

図 4 のように可視化することでレビューの全体像が把握しやすくなると期待される。図 4 では同系色において色が最も濃いものは最も肯定的なグループの文となっており、反対に最も薄いものは最も否定的なグループの文となっている。

6. まとめ

本稿では、文毎にどのくらい肯定的か、どのくらい否定的かという極性の度合とそのカテゴリを自動的に判定するために、SVM を用いた文の極性度合の推定と MIML 学習を用いた文のカテゴリの推定を行った。その結果から、それらの提案手法が各々のベースラインに対し、極性度合の推定について 27%、カテゴリの推定について 21%、精度を向上させる

ことが分かった。今後は SVM によるレビューの極性判定精度が低いことによる極性度合の推定への影響について調査する。また、本研究ではレビューを構成する文のカテゴリを最も可能性の高いものを一つずつしか推定していないが、実際の文が複数の内容を含んでいることに対する対応をどうするのかという課題を解決する必要がある。

参考文献

- [1] Bing Liu, Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, 2012
- [2] Peter D. Turney, Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In Proceedings of ACL, pp.417-424, 2002
- [3] Bo Pang, Lillian Lee and Shivakumar Vaithyanathan, Thumbs up? Sentiment Classification using Machine Learning Techniques., In Proceedings of EMNLP 2002. pp. 79-86, 2002
- [4] 小堀脩豊, 評判分析に基づく推薦システムの研究, 豊田工大, 修士論文, 2013.
- [5] Nguyen.Pham Thanh.Thao, 林貴宏, 尾内理紀夫, 西岡悠平, 竹中孝真, 森正弥, 評価分析における評価対象のカテゴリ分類 - 楽天トラベル「お客様の声」の分析を例として-, DEIM Forum 2009 B6-5, 2009
- [6] Mihai Surdeanu, Julie Tibshirani, Ramesh Nallapati, Christopher D. Manning Multi-instance Multi-label Learning for Relation Extraction, EMNLP-CoNLL '12, 2012