

過賞賛のレビューの検出と特徴語抽出 Detection of Over-Praised Reviews and Feature Word Extraction

草刈祐子^{#1} 中藤哲也^{#2} 鈴木孝彦^{#2} 廣川佐千男^{#2}
Yuko Kusakari Tetsuya Nakatoh Takahiko Suzuki Sachio Hirokawa

1. はじめに

オンラインショップでの商品購入においては、その商品のレビューが購入判断の参考にされることが多々ある。しかし、それらのレビューには有益な情報がなくただ商品を褒めているだけのもの(褒めすぎレビュー)が多く含まれる。そのようなレビューをスパムフィルターのような仕組みで除外できればオンラインショップのユーザーにとって有益である。

あるレビューが「褒めすぎレビュー」であるか否かを教師付き機械学習で判断するためには、褒めすぎレビューの正例が大量に必要である。しかしながら、トレーニングデータを準備することは必ずしも容易ではない。

本稿では、商品に対するレビューの得点分布から、褒めすぎレビューのトレーニングデータを抽出することを試みた。楽天市場の商品レビューデータについて、最高評価の星 5 レビューばかりが大量に投稿されている商品に注目する。そのような商品に対する星 5 レビューを褒めすぎレビューとみなし、SVM を用いた機械学習による判別および特徴語の抽出を試みた。

2. 関連研究

Kokkodis ら[1]は、Amazon レビューのいくつかのジャンルについて、確認済購入者(Verified Purchaser)によるレビューと、それ以外のレビューを比較して、レビューの特徴と有用度を分析した。

林ら[2]は楽天レビューの「ダイエット・健康」「おもちゃ・ホビー・ゲーム」「パソコン・周辺機器」の 3 つのジャンルについて、未購入者のレビューに注目し特徴と有用性を分析した。

草刈ら[3]は、楽天レビューの「ダイエット・サプリ」ジャンルについて、5 点評価のレビューの割合が 8 割以上の商品を収集し、相当する 5 点レビューの特徴を目視で分析した。

本研究ではジャンルを限定せずに、5 点レビューが 7 割以上の商品に対するレビューを褒めすぎレビューととらえ、その特徴を分析することを試みている。

3. 楽天市場データセット

本研究では、楽天株式会社が国立情報研究所の「情報学研究レポジトリ」[4]を通じて研究者に提供している楽天データセット中の、楽天市場商品レビューデータ[5] (以後楽

天レビューデータ)を対象とする。表 1 に使用したデータの項目を示す。

表 1 今回使用したデータの項目

ラベル	説明
商品 ID	商品毎に固有の英数字文字列
商品名	商品の名前 宣伝文句も含まれている
評価点数	1~5 の 5 段階評価。
レビュー本文	主な分析対象
レビュー日時	レビューの投稿された日時
その他の項目	今回は使用しないため省略

4. 褒めすぎレビューの定式化と SVM による学習

本研究では 2010 年 3 月における全レビュー 1,280,239 件のうち、1 つの商品に対する星 5 レビューが 100 件以上、かつその商品の全レビューのうち 7 割以上が星 5 のものを褒め過ぎレビューとし、実験の正例とした。

以下に、使用した褒めすぎレビューの件数内訳を示す。

表 2 対象のレビュー内訳

	レビュー数	商品数	5 点評価割合
正例	18,776	116	0.7 以上

4 で定めた褒めすぎレビューに対し、1 つの商品に対する星 5 レビューが 100 件以上、かつその商品のレビューの星 5 率が一定以下の星 5 レビューを信頼できるレビューと考え負例とした。また、負例の閾値は 3 割以下から 7 割以下まで 0.1 刻みで変化させた。

正例、負例のレビューに現れる語でレビューが機械的に識別可能であるか、そしてこれらのレビューの特徴語を SVM+属性選択[6]を用いて抽出可能であるか実験する。

SVM+属性選択は、正例、負例中のそれぞれの集合に共通する特徴語を抽出するのに有効である。

5. 実験結果

SVM による褒めすぎレビューの識別実験を行った。属性選択なしでの識別性能を表 3、属性数を 1000 とした場合の識別性能を表 4 に示す。

† 1 九州大学 システム情報科学府

Faculty of Information Science and Electrical Engineering, Kyushu University

† 2 九州大学 情報基盤研究開発センター

Research Institute for Information Technology, Kyushu University

表3 属性選択なしの結果

負例閾値	Precision	Recall	F 値	Accuracy
0.3	0.9888	0.869	0.925	0.8679
0.4	0.9548	0.8744	0.9128	0.8703
0.5	0.8562	0.8726	0.8643	0.8561
0.6	0.6869	0.8614	0.7643	0.8066
0.7	0.5479	0.842	0.6638	0.7627

表4 属性数が1000のときの結果

負例閾値	Precision	Recall	F 値	Accuracy
0.3	0.9916	0.9143	0.9514	0.9124
0.4	0.9605	0.9021	0.9303	0.8951
0.5	0.875	0.8704	0.8727	0.8666
0.6	0.7024	0.8656	0.7755	0.8177
0.7	0.5272	0.8715	0.657	0.7469

負例閾値は低い(条件が厳しい)方が抽出の精度は上がり、F値は95%に達した。逆に「褒めすぎレビュー」以外の星5レビューを負例に用いた場合はF値が0.7を切った。

図1に示すように、負例閾値が0.3の場合属性選択の有無はほとんど影響を与えなかった。また、負例閾値が0.4以上の場合は属性(特徴語)数が1000になるまで識別性能が向上した。このことは、正例の識別に必要な単語数が多いことを示唆する。

閾値0.3の場合では、正例ではなく、負例の特徴を識別していると解釈できる。

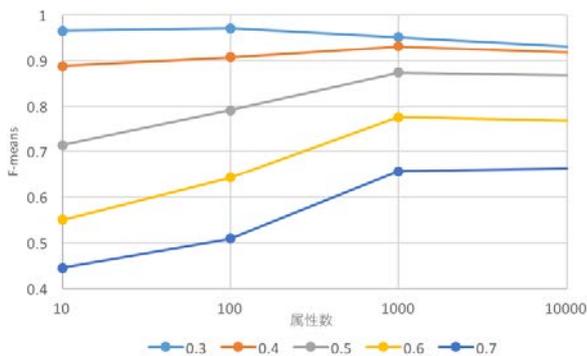


図1. 負例閾値によるF値の違い

抽出された正例、負例の特徴語を表5, 6に示す。

表5 正例の特徴語

閾値	正例の特徴語
0.3	ゲル 受注 タオル 美味しい DVD 大きい 保育園 つば 被り バケツ
0.4	ポスター 被り ダウニー キャス 靴下 名前 ポーチ サンプル 保育園 石
0.5	ポスター 保育園 印鑑 キャス ダウニー 実印 抹茶 ソックス リーバイス 石
0.6	ポスター キャス 抹茶 米ダウニー 石 発売 印鑑 リーバイス バケツ
0.7	ポスター キャス ダウニー 米洗米 嵐 発売 コント レックス 雑穀 お花

表6 負例の特徴語

閾値	負例の特徴語
0.3	禁煙 着タバコ 半熟 ブラシ ネイビー 煙丈巻 く た
0.4	大福 禁煙 ナン 寝心地 麺着 発色 傘前髪 タバコ
0.5	シャンプー 着大福 髪ダイエット 組み立て 本体 禁煙 ナン 寝心地
0.6	ファンシー うどん 海苔 コンタクト ワッフル ピザ ミシン チーズ 髪着
0.7	小瓶 ファンシー ミシン びわこ コンタクト ヒカリ 麺餃子 うどん 着

正例、負例ともに「ゲル」「ポスター」「タバコ」「シャンプー」といった具体的な商品の名称が入っていることがわかる。

SVM+属性選択で褒めすぎレビューを判別できた。しかし、褒めすぎレビューを特徴付けるような少数の特徴語を抽出することはできなかった。さらに特徴語を精査すると、商品のジャンルや種類に関係するものが多かった。

6. まとめと今後の課題

今回はジャンルを区別せずに実験を行ったが、その結果褒めすぎレビューに共通される特徴語よりも、特定の商品やジャンル特有の特徴語が入っていた。このことから、ジャンルを分けて実験を行うことでこれを除くを試みたい。

褒めすぎレビューの抽出については、単純な星5割合以外に、統計的手法の適用[7]も検討している。

参考文献

- [1]Kokkodis, M., Lappas, T., The relationship between disclosing purchase information and reputation systems in electronic markets, ICIS 2016
- [2]林驍, 伊東栄典, 廣川佐千男, 大規模 EC サイトの商品レビュー傾向分析, 情報処理学会研究報告, Vol. 2015-ICS-181, No. 7, 2015
- [3]草刈祐子, 林驍, 鈴木孝彦, 伊東栄典, 廣川佐千男, 有用な商品レビューの特徴と抽出について, 情報処理学会研究報告, 火の国シンポジウム, B3-4, 2018
- [4]NII, 情報学研究リポジトリ, <http://www.nii.ac.jp/dsc/idr/> (accessed at Jan.19.2018).
- [5]NII, 楽天株式会社, 楽天データセット, <http://www.nii.ac.jp/dsc/idr/rakuten/rakuten.html> (accessed at Jan.19.2018).
- [6]Toshihiko Sakai, Sachio Hirokawa, Feature Words that Classify Problem Sentence in Scientific, IIWAS'2012 Special Issue Journal: Intl J of Data Warehousing and Mining (IJWDM), 2012
- [7]中藤哲也, 草刈祐子, 鈴木孝彦, 廣川佐千男, 統計的性質を用いた通販サイトにおけるレビューの評価の試み, 電子情報通信学会技術研究報告, 人工知能と知識処理研究会(AI), ISSN 0913-5685, 2018