

有用な商品レビューの特徴と抽出について

草刈 祐子^{†1} 林 驍^{†1}
鈴木 孝彦^{†2} 伊東 栄典^{†2} 廣川 佐千男^{†2}

概要: オンラインショップでの商品購入においては、その商品レビューを参考にすることが多い。本発表では、購入者の視点から、購入決定のため有用なレビューを抽出する枠組について述べる。レビュー投稿者の属性および各商品あたりのレビュー回数と評価ランクの分布に注目し、「役に立たないレビュー」を除去することで有用なレビューを抽出することを目指す。

キーワード: テキストマイニング, 機械学習・データマイニング

Extraction Methods of Useful Product Reviews with Feature Selection

Yuko Kusakari^{†1} Xiao Lin^{†1}
Takahiko Suzuki^{†2} Eisuke Ito^{†2} Sachio Hirokawa^{†2}

Abstract: Customer reviews are important reference materials for online shopping customers. In this article, we discuss two methods of filtering out useless product reviews from the viewpoint of customers. We focus on attributes of reviewers, the number of reviews per product, and the average score and the count of each product of reviews. Useless reviews can be filtered out by using those features.

Keywords: text mining, machine learning, data mining

1. はじめに

近年ネットでの買い物が普及してきた[1]。ネットショッピングは便利ではあるが、直接手にとって選ぶことができないため店舗で購入するより情報が乏しい。そこで口コミやレビューは、同じ消費者の立場からの意見が聞けるため購入を決める大きな手助けになる。欲しいものの中から評価が高いものを選んで買うこともあるだろうし、買おうと思ったものの評価が低かったため購入を見送ることもあるだろう。ネットショッピングにおいてレビューの与える影響は大きい。しかし、中には役に立たないレビューもある。明らかに正しくないと思えるもの[2]もあれば、購入に必要な情報が含まれていないもある。本研究では、レビューを役に立たないレビューの特徴を抽出・除去することによって役に立つレビューを抽出する手法について論じる。

本研究で用いた、特徴語を抽出し文書を分類する手法については足立ら[9]を参考とした。レビューの機械学習については山根ら[10]など、多くの研究がある。

本論文の構成は以下の通りである。第2章では使用するデータについて述べる。第3章では未購入者のレビューの特徴を学習することによって役に立たないレビューの抽出を試みる。第4章では大量のレビューが投稿される商品のレビューの特徴を調べる。最後に第5章でまとめと今後の課題を述べる。

2. 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」[3]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[4]（以後楽天レビューデータ）を対象とする。表1に本研究で利用した楽天レビューデータの主な構造を示す。

^{†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 1. 楽天レビューデータの主な構造

ラベル	説明
投稿者	匿名化 ID. 投稿時匿名の場合は「購入者さん」
商品 ID	商品毎に固有の英数字文字列
商品名	商品の名前 特長など
ショップ名	ショップの名前
ジャンル ID	ジャンル毎に固有の数値
評価点数	1~5 の 5 段階評価.
レビュータイトル	レビューのタイトル.
レビュー本文	レビューの中身. 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	今回は使用しないため省略

3. 未購入者レビューの特徴によるフィルター

林ら[5]は, 楽天レビューデータ中に, 未購入者のレビューが含まれていることに注目した. 彼らは未購入者レビューを, 商品の購入とは無関係な「役に立たないレビュー」の学習正例として使えると考えた. 未購入者のレビューの特徴語を抽出し, 役に立たないレビュー一般を除去するために利用することを提案している.

林らは商品の評点スコアが高評価(5 点)である未購入者のレビューを正例, 評価 5 点の購入者レビューを負例とし SVM で学習し, 特徴語を抽出している.

林らの手法に基づき, Health, Game, Computer の 3 ジャンルの楽天レビューデータについて, MeCab でレビュー文中の単語を切り出し[6]属性選択を行う線形 SVM[7]を用いて未購入者の 5 点評価レビューを学習した後, テストデータに対する正答率(accuracy)を 5 分割交差検定で評価した. その結果を表 2 に示す.

表 2. 未購入者の 5 点評価レビューの正答率

ジャンル	Max. accuracy	特徴語数
Health	0.9518	16
Game	0.8839	40
Computer	0.9520	60

比較的小数の特徴語によって, 未購入者の 5 点評価レビュー(役に立たないレビュー)が判別可能である. Health ジャンルについて, 特徴語の一覧を表 3 に示す.

林らは, 未購入者の低評価レビュー(評点 1)についても「役に立たない」レビューの学習正例として利用している. 高評価の場合と同様に評価 1 点の購入者レビューを負例と

して SVM で学習し評価した結果を表 4 に示す.

表 3. 未購入者高評価レビューの正負特徴語

順位	頻度	正の特徴語	頻度	負の特徴語
1	1928	助詞: ので	2323	助詞: は
2	1032	記号: !	2331	助詞: と
3	1813	名 詞: 楽し	2703	名詞: 円
4	1232	助動詞: な	2842	名詞: リピ
5	1535	名詞: 購入	5018	名詞: 効果
6	977	動詞: 届く	2596	形容詞: 美味し
7	2790	名詞: 今	6358	助動詞: で
8	2307	記号: (	6047	動詞: いる

表 4. 未購入者の 1 点評価レビューの正答率

ジャンル	Max. accuracy	特徴語数
Health	0.6828	100
Game	0.6057	40
Computer	0.7167	20

3.1 未購入者レビューフィルターの評価

未購入者によるレビューが, 一般の「役に立たないレビュー」の判別に有効か否かを評価するために以下の実験を行った.

- 対象となるジャンルの, 購入者による全レビュー(Health: 2, 312, 869 件, Game: 1, 465, 269 件, Computer: 1, 261, 171 件)を入力とする.
- それぞれのジャンルの未購入者高評価・低評価レビューで学習した SVM 3×2 種類で購入者のレビューを判別する.
- それぞれの SVM 出力において, スコアが高いもの 100 レビューを選択する.
- 選択されたレビューそれぞれに対し, 次の各項目に当てはまるか否か人間がチェックする.
 - 商品に対するものではない
 - 商品を使用していない
 - 情報が薄い
 - 参考にならない

人間によるチェックの結果を図 1 に示す. 未購入者の 5 点評価レビューの特徴語から作成した「役に立たないレビュー」フィルターが有効に機能し, 「情報が薄い」や「参考にならない」レビューを抽出している(recall 0.75~0.99).

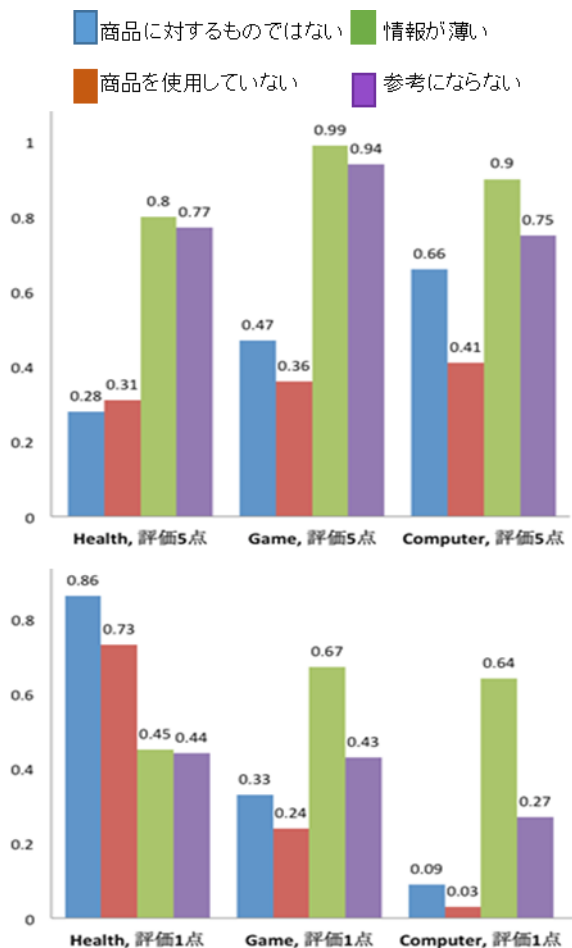


図 1. 未購入者レビューフィルターの人間による評価 (recall)

3.2 未購入者レビューによる手法の限界

未購入者レビューを用いて、役に立たないレビューの特徴を抽出する手法には、次に挙げる限界がある。

(1) 汎用性の問題

楽天レビューデータと異なり、多くの EC サイトでは、未購入者が商品レビューを投稿することを禁止している。そのようなサイトのレビューに対して本手法は適用できない。

(2) 利用可能なレビュー数

楽天レビューデータにおいても、未購入者レビューはユーザーレビュー全体の 0.2~0.3% に留まり、サンプルとして利用可能なレビュー数が少ない。

(3) レビューの内容

未購入者高評価レビューには、注文した商品が届く前に投稿された「楽しみにである」という内容が含まれる (表 3 正の特徴語参照)。これらのレビューは、確かに購入の判断の役に立たないが、特徴語抽出のためにふさわしいか否か疑問が残る。

(4) 低評価レビューの扱い

未購入者低評価レビューを「役に立たないレビュー」と

みなし、SVM 学習によって判別した結果の性能は必ずしも高くない。低評価レビューの扱いを考えなおすべきである。

これらの限界を解決するために、「役に立たないレビュー」の学習正例獲得について、4 章に示す手法を提案する。

4. 商品毎のレビュー集合に注目した手法

先行研究[8]では、レビュー評価点の分布は正規分布になるとされている。しかしながら、楽天レビューデータにおける評価点の分布は図 2 に示すように、評価点 5 が非常に多く、不自然である。また、3 章の結果から、評価点 5 の未購入者レビューを学習正例とすることで、役に立たないレビューの検出がうまくいくことがわかった。

本章では高評価のレビュー全体から「役に立たないレビュー」の学習正例を抽出するため有用な一手法を提案する。

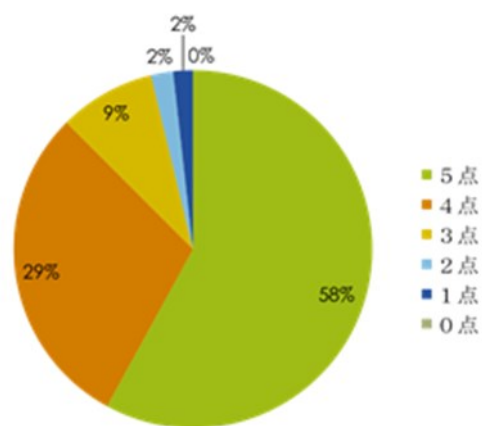


図 2. 楽天レビューデータにおける評価点毎のレビュー数の割合

4.1 高評価レビューの分類

高評価のレビューは次のように分類できる[6]。

- A) 配送の迅速さなどショップの対応についてのレビュー
- B) 価格が安いなどの入手コストについてのレビュー
- C) 入手前の期待など客観的評価に基づかないレビュー
- D) 製品を評価するレビュー

A), B), C) については、製品を選択する上では「役に立たない」レビューと考えられる (ショップ選択においては有用な場合もある)。A), B), C) のレビューの例を人間が抽出し、特徴を機械学習することは実現可能と思われる。

D) に属するレビューから「役に立たない」レビューを選択することを考える。本論文では一商品に対するレビュー集合の評価の分布に注目して選択を試みる。

4.2 レビューが大量に投稿される高評価商品

ある商品に対して大量のレビューが投稿され、かつ評価の分布が高評価に偏っている理由は次のようなものが考えられる。

- i. 4.1 節 A), B), C)のレビューに分類されるレビューが大部分を占めている。
- ii. 該当商品が本当に良い
- iii. いわゆるやらせレビューが多い
- iv. 商品またはショップの熱狂的ファンによる主観的な評価

i, iii, iv については「役に立たない」レビューとみなすことができる。ii に属するレビューは「役に立つ」レビューである。しかしながら、大量に高評価レビューが投稿されるような良い商品は、商品レビュー以外の場でも良い評判があるはずであり、他の手段で特定が可能であると思われる。

以下では、『「大量のレビューが投稿されかつ評価の分布が高評価に偏っている商品」』のレビュー（以下「褒めてばかりのレビュー」）のサンプルを抽出した結果を示す。

楽天市場全体では1ヶ月に約130万件のレビューがあるため、役に立たないレビューが多いと思われるダイエット・健康ジャンルのうち2010年1月分で1000件近くのレビューのあった2つのジャンルのレビューについて調査した。表5に内訳を示す。

表5. 使用したジャンルのレビューデータの内訳

ジャンル	全体の レビュー数	レビュー が5件以上
ダイエット・サプリ	170 商品 /1243 件	23 商品 /652 件
ダイエット シェイク >>セット	54 商品 /958 件	21 商品 /878 件

この2つのジャンルに含まれる商品の中から2010年1月の1ヶ月間に5件以上のレビューを獲得している商品を抽出し、該当商品に対する評価点5のレビューがどのような傾向を示しているか分類した。

まず、配達物の速さやショップの対応を評価されている商品(4.1 分類A)と安さを評価されている商品(4.1 分類B)を目視で分類した。表6,表7に分類A), B)のレビューの例を示す。

次に、8割以上のレビューが評価点5である商品を抽出した。これらの商品に対するレビュー（褒めてばかりレビュー）の例を表8に示す。目視でチェックしたところ、レビュー中に誇大と思われる賛辞がしばしば見受けられた。

表6. ショップの対応を評価するレビュー(分類A)の例

素早い対応、有難うございます。頑張ってみたいと思います。
注文したその日に発送され、迅速な対応がよかった。
大変おそくなりました。迅速な対応ありがとうございました。今後ともよろしくお願いします。

表7. 安さを評価しているレビュー(分類B)の例

気になる新しいサプリを100円で試せてとてもありがたいです。
安くよかった。次回もぜひ利用したいと思いました。
いつも気になっていたショップで、今回は送料無料だったので初購入してみました。届くのが楽しみです。良かったらまたリピしたいと思います。

表8.褒めてばかりレビュー（評価点5が8割以上）の例

慢性便秘だけど、薬は飲みたくない私。お腹にたくさん貯めないことが、ダイエットをする上での重要な要素になってきます。夜飲めば朝は必ず出ますよ。
2人で飲むと、あっと言う間になくなり、(あたり前ですが・・・)リピート購入です。痩せるという実感は、今のところありませんが、翌朝の効果は・・・◎ 体の調子もいいので、お気に入りの商品です。
これ、いいです！夕食後に6錠飲んで、翌日、朝食後にスッキリ。しかもおなかが痛くならない!!驚きました。

なお、本実験では、表6,7に対応する商品はすべて、評価点5のレビューが8割未満であった。

商品当たりのレビュー数が5以上でかつ、分類A), B), 「褒めてばかりのレビュー」いずれにも当てはまらない商品のレビュー（評価点5が8割以下であるが高評価）の例を表9に示す。

表9. その他の高評価レビューの例

これを飲むようになってから翌日目が腫れていることがあまりないので手放せません。私は夜一錠で少しむくみがある時は翌朝一錠飲んでみます。
立ち仕事&むくみ体質の私には手放せないサプリ。もう半年ほど愛用しています。飲みやすい小さめのカプセルもポイント高し。
(前略)しかし去年の冬に飲みだしてから毎日欠かさず飲んでいたので良かったのか 今季は冬入った頃の寒さが激しくなった頃は、少ししもやけが出ましたが、その後もしもやけがでなくなりました！！もう一最高に嬉しいーッ＼(^-^)/ (後略)

5. まとめと今後の課題

5.1 まとめ

本稿では、有用な商品レビューを抽出することを目的とし、「役に立たないレビュー」の例を収集し、それらを除去するための手法を提案した。

3 章では、楽天レビューデータから「未購入者のレビュー」を抜き出し、SVM で学習する手法を使って「役に立たないレビュー」を検出することを試みた。未購入者の 5 点評価レビューは、役に立たないレビューの学習正例として有用であることを示した。

次に、未購入者のレビューの例は入手困難な場合があるため、一般的に入手可能な学習例を収集することを考えた。

4 章では、「大量のレビューが投稿されかつ評価の分布が高評価に偏っている商品」のレビュー（褒めてばかりのレビュー）に注目した。褒めてばかりのレビューと、他のレビューのサンプルを試験的に収集し、目視で評価したところ、褒めてばかりのレビューが、役に立たないレビューの学習例として有望であることが分かった。

5.2 今後の課題

褒めてばかりのレビューを正例として学習し、役に立たないレビューを実際に判別することが当面の課題である。

褒めてばかりのレビューの中に「役に立つレビュー」(4.2 節の ii)が含まれているか、含まれている場合は自動的に判別可能か確認する必要がある。

学習にあたっては、他の種類の「役に立たないレビュー」(4.1 節の A), B), C)) と「褒めてばかりのレビュー」の学習を独立して行う可否かを考慮する。

最後に、本稿では主に高評価の役に立たないレビューについて扱ったが、低評価のレビューについても、役に立た

ものを除去する方法を別途検討する。

参考文献

- [1]経済産業省商務情報政策局情報経済課，平成 28 年度我が国経済社会の情報化・サービス化に係る基盤整備(電子商取引に関する市場調査)報告書. 2017.
- [2]J-Cast ニュース，「やらせ」のロコミや架空の注文なんと 11 万件超! 「楽天市場」で大掛かりなステマが発覚，<http://www.j-cast.com/2015/03/21231013.html> (accessed at Jan.19.2018).
- [3]NII，情報学研究リポジトリ，<http://www.nii.ac.jp/dsc/idr/> (accessed at Jan.19.2018).
- [4]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>.
- [5]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
- [6]Sakai, T. and Hirokawa, S. Feature words that classify problem sentence in scientific article, Proc. iiWAS 2012, pp.360-367, 2012
- [7]形態素解析エンジン MeCab, <http://taku910.github.io/mecab/> (accessed at Oct.8.2015).
- [8]Shardan, U. and Maes, P., Social information filtering: algorithms for automating “word of mouth”, CHI '95 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems pp.210-217 ACM Press/Addison-Wesley Publishing Co. New York, NY, USA, 1995
- [9]足立 裕介，鬼村 直弥，山下 貴範，廣川 佐千男 標準的尺度と SVM 尺度を使った属性選択による文書分類の性能評価 (言語理解とコミュニケーション) 電子情報通信学会技術研究報告. A・P, アンテナ・伝播 ISSN: 0913-5685 2016 巻: 116 号: 213 ページ: 1-6
- [10]山根宏彰，萩原将文，Web 上のロコミを用いたキャッチコピー選択システム 2015/04/01, Online ISSN, Print ISSN, https://doi.org/10.14864/fss.30.0_654, https://www.jstage.jst.go.jp/article/fss/30/0/30_654/_article/-char/ja