

化粧品レビューサイトにおけるクチコミの有用性判定

高島侑里 青野雅樹
 豊橋技術科学大学 情報・知能工学課程
 yuritaka@kde.cs.tut.ac.jp, aono@cs.tut.ac.jp

1. はじめに

インターネット上のレビューサイトの普及により、消費者の意見発信が盛んに行われている。企業が発信する情報とは異なり、消費者がレビューサイトに投稿する意見や評判は主観的な情報が多数含まれているという特徴を持つ。そのためある商品やサービスを購入・利用する際に、生の声が聞けるという点で、レビューサイトに投稿された情報を活用する消費者が増加している。例えば、飲食店に対するクチコミが投稿できる「食べログ」では月間約 17.8 億 PV[1]、化粧品に関するクチコミが投稿できる「@cosme」では月間約 2.8 億 PV[2]を記録しており、消費者のレビューサイトへの関心の高さが伺える。

一方、レビューサイトにはユーザが自由に意見を投稿できるため、レビューサイトのクチコミ数は膨大となる。これら全てのクチコミを閲覧し、商品やサービスを選択することはユーザにとって大きな負担となる。またクチコミの質のばらつきも激しく、非常に有益な情報を発信するユーザもいる一方、具体性に欠けるクチコミを投稿するユーザも多い。そのため膨大なクチコミデータベースから質の良い有用なクチコミを抽出して表示できれば、ユーザは効率よく商品やサービスの情報を取得でき、レビューサイトの閲覧時間や労力の面で負担を軽減することが可能である。

本研究では、化粧品レビューサイトのひとつである「@cosme」に投稿されたクチコミの有用性を判定する手法を提案する。@cosme では、クチコミが参考になった場合にそのクチコミに「Like」を投票できるシステムを導入している。本研究では Like の数をクチコミの有用度と定義する。そしてクチコミの構造解析や意味解析などを行うことで素性を抽出し、それらを説明変数とする回帰モデルを構築する。構築した回帰モデルを用い、クチコミの Like の数を推定することを目的とする。

2. 関連研究

近年、商品やサービスのレビューに関する研究は盛んに行われている。Jyoti らは Amazon.in に投稿されたレビューの “helpfulness” を定義し、機械学習

による推定を行い、レビューの読みやすさ、極性情報、主観性、エントロピーなどのレビュー文の質を問う意味的な素性が、helpfulness を推定するうえで重要であるという見解を示している[3]。Iris らはワインレビューサイトに投稿されたレビュー文から、ワインの色、原産国などワインの特徴を推定する研究を行っている[4]。彼らは BoW (Bag of Words)の他に LDA (Latent Dirichlet Allocation)[5] や Word2Vec[6]を用いた意味的な素性を提案し、SVM で分類を行っている。その結果、BoW のみより、LDA や Word2Vec を組み合わせた方が高い F 値でワインの特徴を推定できることを述べている。

@cosme のクチコミデータを用いた研究も広く行われている。松波らは@cosme を閲覧するユーザに対し、真に参考となるクチコミを提示するシステムの構築例を示している[7]。対象ユーザと使用感や嗜好感が類似する投稿者を検索し、その投稿者が発信したクチコミをユーザに提示することで、ユーザが真に参考できるクチコミを抽出できると提案している。使用感や嗜好感が類似する投稿者を推定するために、化粧品の評価項目ごとのレイティングに着目し、ポジティブ・ネガティブ表現辞書を構築してクチコミを評価項目ごとに自動レイティングする手法を実装した。安部らは化粧品の情報収集や、新しい化粧品の検索など、ユーザの様々な目的に応じたクチコミの推薦システムの構築例を示している[8]。ユーザの興味に基づいたクチコミ推薦を行うためにクチコミの評価視点に着目し、人手で評価視点の抽出を行っている。

3. 提案手法

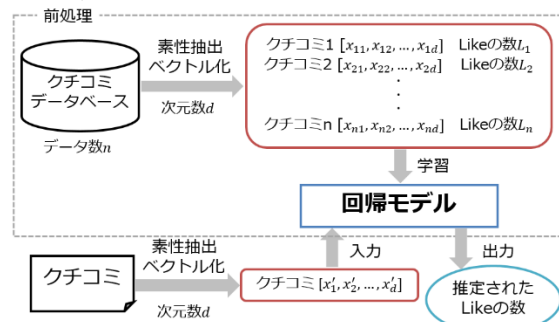


図1 クチコミの有用性判定のシステム図

本研究で構築したクチコミの有用性判定システムを図1に示す。前処理（訓練）として有用度を判定するための回帰モデルの学習を行う。まず、@cosmeに投稿されたクチコミを収集し、クチコミデータベースを作成する。次にクチコミの構造解析、統語的解析、意味解析を行うことで素性を抽出し、クチコミをベクトルに変換する。ベクトル化されたクチコミと、正解ラベルとなる Like の数を回帰モデルに学習させる。

未知のクチコミデータに対しては、訓練時と同様に素性抽出とベクトル化を行い、訓練済みの回帰モデルに入力する。その結果、推定された Like の数をそのクチコミの有用度とする。

4. 提案素性

提案素性は、クチコミの長さや改行の数などの構造的な素性、クチコミの品詞情報を用いた統語的な素性、話の転換やトピックなどのクチコミの質を測る意味的な素性の3つに大別され、全体で、22次元で表現した。具体的には表1に示す通りである。

表1 クチコミの有用度判定に用いる素性

カテゴリ	番号	素性	次元数
構造的な素性	1	クチコミの文字数	1
	2	改行の数	1
	3	星の数	1
	4	購入品/モニター	1
	5	現品/サンプル	1
	6	リピートの有無	1
	7	画像の数	1
	8	感嘆符の数	1
統語的な素性	9	名詞の数	1
	10	動詞の数	1
	11	形容詞の数	1
	12	副詞の数	1
	13	数字の数	1
	14	専門用語の数	1
	15	体言止めの文の数	1
意味的な素性	16	極性	1
	17	話題の転換数	1
	18	話題の詳細度	1
	19	LDA	4

4.1. 構造的な素性

構造的な素性として表1に示す8種類の素性を提案する。最初の3つの素性はそれぞれクチコミ本文から計算を行う。また、クチコミのメタ情報から星の数、購入品/モニター、現品/サンプル、リピートの有無、画像の数を抽出する。

4.2. 統語的な素性

統語的な素性として表1に示す7種類の素性を定義する。このうち、名詞の数、動詞の数、形容詞の数、副詞の数、数字の数、体言止めの文の数は、クチコミ本文を形態素解析システムのMeCab[9]を用いて形態素解析を行い、各形態素の品詞情報から計算する。

専門用語の数はJuman++[10]を用いて抽出を行う。具体的には形態素解析を行った際、Wikipediaから獲得した語彙、未定義語を「専門用語」と定義し、クチコミ内に含まれる専門用語の数を素性とする。専門用語の数が多いほど、化粧品に含まれる成分の具体的な記述や説明、他の化粧品ブランドの製品との比較など、読者に有益な情報がクチコミに含まれていると判断できる。

4.3. 意味的な素性

意味的な素性として極性を用いた素性、話題の転換数、話題の詳細度、LDAを用いた素性の4種類7次元の素性を提案する。クチコミの極性情報、分散表現、潜在的なトピック情報からクチコミの質の解析を行うことを狙いとする。

極性を用いた素性に関して、クチコミに対して式1から求めたpolarity値を素性として用いる。

$$\text{polarity} = s(\text{positive}) - s(\text{negative}) \quad (1)$$

ただし $s(\text{positive})$ はポジティブと判定された文の数、 $s(\text{negative})$ はネガティブの文の数とする。

$s(\text{positive}), s(\text{negative})$ を算出するためには、クチコミを文に分割した後、各々の文に対して極性を付与する必要がある。各々の文に極性を付与する処理に関して、教師あり学習を用いた分類モデルのひとつであるSVM(Support Vector Machine)を用いた極性分類サブシステムを構築することで実現する。

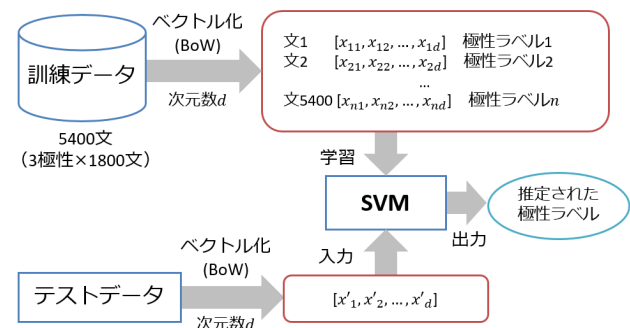


図2 極性分類サブシステム

本研究で構築した極性分類サブシステムを図2に示す。本サブシステムでは、文をポジティブ、ニュー

トラル, ネガティブの3つの極性に分類する. 訓練部では訓練時に必要なコーパスを作成するため, @cosme に投稿されたクチコミから各極性について1800文, 合計5400文の収集を手作業で行い, 極性ラベルを付与している. 次に, 収集した文をBoWでベクトル表現に変換する. BoWの重みづけは, TF-IDFモデルを採用している. その後, ベクトル化したデータを入力データとしてSVMで学習し, 訓練モデルの構築を行う. 極性判定部では, 極性を判定したい文を同様の手法でベクトル変換し, 訓練モデルに入力データとして与える. そして, モデルが出力した予測ラベルをその文の極性とみなす.

話題の転換数, 話題の詳細度はクチコミの各文をベクトル表現に変換し, 隣接する文のコサイン類似度を基に算出する. 話題の転換数は, クチコミがどのくらい多様な観点について記述しているかを示す尺度となる. 例えば化粧水では, うるおい, 美白効果などの評価項目が考えられ, 一般的には多くの評価項目について言及されているほど, 読者はそのアイテムの情報を獲得でき, 有用であるといえる. 一方, 話題の詳細度は, あるトピックについてどの程度詳細に記述されているかを示す. 前述のようにあるアイテムに対する評価項目は複数考えられるが, ひとつひとつの評価項目について綿密に書かれているほど, 読者にその項目に関する情報を多く提供しているといえる.

上述の話題の転換数, 話題の詳細度の算出方法について説明する. はじめに, 話題の転換数, 話題の詳細度を求めるために, クチコミを文単位に分割する. その後, 文をMeCabで形態素解析し, 各形態素をベクトル表現に変換する. このため分散表現の代表例であるWord2Vecを用いる. 化粧品に特化したモデルを構築するため, ここで用いるWord2Vecのコーパスとして@cosmeのクチコミを選択し, 次元数は100としている.

文のベクトル表現は, 各形態素ベクトルの和と定義する. クチコミを構成する全ての文をベクトル表現に変換した後, 隣接する文同士のコサイン類似度を算出する. コサイン類似度が閾値以下の部分を話の切れ目, すなわち話題が転換したと判断し, 閾値以上であれば話題が連続しているとみなす. 素性の一つである話題の転換数convはコサイン類似度が閾値以下となった回数と定義する. 一方, 話題の詳細度detailは式2で算出する. ただしnはクチコミの文の数とする.

$$detail = \frac{n - conv - 1}{conv + 1} \quad (2)$$

潜在的なトピック推定手法のひとつであるLDAを用いた素性ではクチコミ内の素性単語の数, 単語適合率, 単語再現率, 単語出現確率の総和の4つを導入する. LDAを用いて, 多様な評価視点やトピックを考慮した素性の抽出を試みる.

はじめに, 化粧品に特化したLDAのトピックモデルを構築するため, コーパスとして@cosmeに投稿されたクチコミを使用する. 次に, コーパスを用いて構築したモデルからトピック群を得る. 本研究ではトピック数を100に設定した. その際, 各トピック群について, そのトピックに出現する単語と, その出現確率を計算する. 次に構築したLDAモデルを用い, 各クチコミが, 得られたトピックを潜在的に有する指標としてトピック確率を計算する. 推定確率の最も高いトピックに属する単語群を5件抽出し, そのクチコミの素性単語とする. 最後に, 抽出した素性単語からクチコミ内の素性単語の数, クチコミの単語適合率, 単語再現率, 単語出現確率の総和の4つの素性値を算出する. 単語適合率(term precision), 単語再現率(term recall)の素性値算出式を式3, 4に示す. なお, cはクチコミ内の単語と素性単語との適合数(クチコミ内の素性単語の数), nは素性単語数, dはクチコミ内の単語数である.

$$\text{term precision} = \frac{c}{n} \quad (3)$$

$$\text{term recall} = \frac{c}{d} \quad (4)$$

クチコミ内の単語と素性単語で適合した単語の出現確率の総和を式5で表す. なお, w_j はクチコミ内の単語と素性単語で適合した単語を示す.

$$\sum_{j=1}^c \text{probability}_{w_j} \quad (5)$$

5. 評価実験

提案手法の有効性を確認するため評価実験を行った. 提案素性を用いたクチコミの有用性判定についてベースラインと比較した評価を行う.

5.1. データセット

@cosmeからhtml文書を収集しデータの抽出を行った結果, 88455件のクチコミデータを取得した. 本実験ではある一定の期間を過ぎ, Like数がほぼ収束しているクチコミとしてLikeの数が5以上, 100以下のクチコミ16713件を用いることとする.

5.2. 評価尺度

クチコミの有用性判定システムの評価尺度として

R^2 (寄与率) と RMSE (Root Mean Square Error, 平均二乗誤差の平方根) を用いた。寄与率は標本から求めた回帰式の当てはまりの良さを示すものである。RMSE は予測値が正解値からどれだけ乖離しているかを示す評価尺度である。

5.3 実験方法

本実験ではクチコミデータセットを 5 交差検定で検証し、回帰モデルとして SVR(Support Vector Regression)を使用した。カーネルは線形カーネルとした。ベースラインに使用する素性として BoW を選択し、TF で重みづけを行った。

5.4 実験結果

クチコミの有用性判定システムの実験結果を表 2 に示す。

表 2 クチコミの有用性判定の実験結果

手法	次元数	R^2	RMSE
ベースライン	14149	0.0910	0.1130
提案手法	22	0.1094	0.1119

5.5 考察

本節ではクチコミの有用性判定システムについての考察を行う。実験結果より、全ての評価尺度において提案手法がベースラインを上回っていることを確認し、提案手法の有効性を示した。

また、ベースラインと提案手法を統合し、Random Forest の重要度で素性選択を行った場合についても精度を比較した。更に最も精度が良かった次元数について、グリッドサーチを行い回帰モデルのパラメータ調整を行った。これらの実験結果を表 3 に示す。

表 3 素性選択とグリッドサーチの実験結果

手法	次元数	R^2	RMSE
連結素性+素性選択	700	0.1584	0.1088
連結素性+素性選択 (グリッドサーチ後)	700	0.1701	0.1079

表 3 より、次元数が 700 のとき最も優れた評価値が得られた。素性重要度による素性選択を行い、ノイズとなる素性を除去することで大幅な精度向上を確認した。また、グリッドサーチによる回帰モデルのパラメータ調整により、寄与率、RMSE のいずれの評価値においても更なる精度向上が確認できた。

6. おわりに

本研究では化粧品レビューサイトに投稿されたクチコミの有用性を判定する手法を提案した。提案手法としてクチコミの構造的情報や統語的情報だけでなく、Word2Vec による分散表現や LDA を用いたトピック推定を応用し、クチコミの質を解析する意味的な素性を導入した。その結果、ベースラインである BoW より優れた評価値が得られることを示した。また、ベースラインと提案手法を組み合わせ、素性選択を行うことで更なる精度向上を実現した。

今後の課題としては、@cosme 以外の他のレビューサイトのデータを提案手法に適用した場合の有効性の検証が挙げられる。また、本研究では評価実験に用いるデータとして、Like の数が一定数以上付与されているクチコミを選択した。本研究の最終的な課題として、Like の数が付与されていない未知のクチコミデータに対する本システムの妥当性の評価が挙げられる。この課題を実現するためには、あるアイテムについてのクチコミの作成と、正解データとなる有用性を表す指標を客観的に付与する必要があり、被験者による評価実験が必要となる。

謝辞

本研究の一部は、科研費基盤 (B) (課題番号 26280038) の支援を受けて遂行した。

参考文献

- [1] <http://user-help.tabelog.com/advertisement/>
- [2] <http://www.istyle.co.jp/business/uploads/sitedata.pdf>
- [3] Jyoti P.Singh, Seda Irani, et al., "Predicting the "helpfulness" of online consumer reviews", *Journal of Bus. Res.* 70, pp.346-355, 2017.
- [4] Iris Hendricks, Els Lefever, et al., "Very quaffable and great fun: Applying NLP to wine reviews", *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp.306-312, 2016.
- [5] David M. Blei et al, "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, 3, pp.993-1022, 2003
- [6] <https://code.google.com/archive/p/word2vec/>
- [7] 松波友稀, 上田真由美, 中島伸介, 階上猛, 岩崎素直, John O'Donovan, Byungkyu Kang, "コスメアイテム評価表現辞書を用いた評価項目別レビュー自動スコアリング方式", DEIM2016, 2016.
- [8] 安部小百合, 小林一郎, "ユーザの属性に基づく化粧品レビュー文書の推薦", 言語処理学会 第 22 回年次大会 発表論文集, pp.147-149, 2016.
- [9] <http://taku910.github.io/mecab/>
- [10] <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN++>