

研究指導 中澤 真 准教授

グルメサイトにおける有用な口コミ文の抽出アルゴリズム ～単語の出現率に着目して～

遠藤 美香子

1. はじめに

近年、国内グルメサイトは「食べログ」や「ぐるなび」といった大手4社がしのぎを削っている[1]。各社における競争に伴い、グルメサイトのサービスは向上してきている。これにより、グルメサイトが利用しやすくなり、外食するときの店を探す際に欠かせないものになりつつある。

グルメサイトには店の場所や営業時間などの基本情報が掲載されているが、それ以上に大きな役割を持つのが実際に店を利用した人によって投稿された口コミ情報である。口コミには、店の雰囲気や食べたものの感想、混み具合、客層など、基本情報では得ることのできない消費者の生の声が書かれている。このため、多くのユーザが口コミを参考にして店を決めている[2]。

一方で、グルメサイトの口コミは情報量が多く、求める情報をすぐに見つけられないという問題がある。これに対して、冗長な口コミをフィルタリングして有用な情報を効率的に抽出しようという研究がなされている。しかし、これらの多くはユーザのニーズに合わせた口コミを選び出す試みであり、誰にとっても有用な情報だけを取り出すには適していない。

そこで本研究では、単語の出現頻度によって口コミの文単位で要不要を判断する方法を提案し、多くのユーザが求める有用性の高い文のみを抽出することを目指す。

2. グルメサイトにおける口コミの現状と課題

2.1 口コミの現状

グルメサイトにおける口コミは、ユーザが店を決める上で重要な役割を果たしている。MMDLabo株式会社の調査では[2]、グルメサイト・アプリの口コミを参考にして店に行く人は約7割を超えているという結果が報告されている。

グルメサイトのシェア1位が「食べログ」である。同業他社である「ぐるなび」、「ホットペッパーグルメ」といったサイトと比べてみても、アクセスシェアで食べログは55%と最も高くなっている[3]。また、掲載店舗数や口コミ数は食べログが最も多く、なかでも口コミ数は2000万件を超えており、国内のグルメサイトの中でも圧倒的な数を誇っている[4]。このデータからも、クチコミサイトが一

般生活の中に浸透して、飲食店を選ぶ際に口コミが重要な役割を果たしていることは明らかである。

2.2 グルメサイトにおける口コミの課題

広く利用されているグルメサイトの口コミであるが、情報量が多すぎて必要な情報を取得しづらいという問題点もある。特にグルメサイトの口コミは、オンラインショッピングサイトなどの口コミに比べて投稿された口コミの文章が長く冗長な傾向にあり、求める情報を探し出すことをより困難にしている。

実際に食べログの口コミと、オンラインショッピングサイトの「Amazon」の口コミの文字数を比べてみると、Amazonの口コミでは短いものと50文字未満のものも多くあるのに対して、食べログの口コミは短いものでも200文字を超えるものがほとんどである。

このことからグルメサイトの口コミは他の口コミサイトと比較して長い傾向にあることがわかる。これはグルメサイトの口コミの中には、「仕事のついでに寄りました」「友人が誕生日だったため～」などの飲食店利用のシチュエーションの説明など、店舗やそのサービス内容と直接関係のない記述が多く含まれているためだと考えられる。

このため、冗長な口コミをフィルタリングして有用な情報を効率的に抽出しようという研究がなされている。

3. 口コミ文の分析

口コミ文のフィルタリングに関する研究はいくつか行われている。毛ら[5]は膨大な数の口コミからユーザに役立つ情報を提示するため、人それぞれの価値観に基づいて口コミをフィルタリングする方法を提案している。ここで述べている価値観とは、「何が大事で何が大事でないか」という判断のことである。この研究では評価属性に言及した口コミを抽出している。映画であれば「ストーリー」や「俳優」、デジカメであれば「デザイン」、「操作性」といった商品特有の評価属性があり、ユーザがこだわりを持つ評価属性の語句が含まれているか否かで口コミをフィルタリングしている。

桑田[6]らの研究も口コミを評価属性によって分類している点では毛らの研究と共通点が多いが、共起を踏まえて分類している点で異なっている。

これらの研究はいずれもオンラインショッピングサイトを対象に、ユーザごとに異なる価値観や視点に適した口コミを抽出することを目的としている。しかし、グルメサイトではユーザが求める情報(評価属性)は、「料理の味」や「店の雰囲気」などのようにユーザ全体で共通であるものが多い。加えて、グルメサイトの口コミは長い文章が多いため、1件の口コミの中で有用な文と不要な文をフィルタリングし、有用性が高い文のみを提示することが効果的だと考える。

そのため本研究では有用性の高い文の基準を定義し、口コミの文単位で要不要を判断して、必要な文のみを抽出することを目指す。

4. 分析方法と結果

4.1 人手による必要な文と不要な文の判別

本研究ではグルメサイトの中で最も口コミ数の多い「食べログ¹」のデータを対象に分析を試みる。利用するデータとして、食べログに掲載されている口コミをランダムサンプリングで80件選び出し、その内50件を要不要の判断基準を算出するためのデータとして利用し、残り30件を検証用データとして用いることにする。

まず、口コミを必要な文と不要な文に人手によって分類するための基準を定義する。客層や接客態度など、ユーザが店を選ぶ際に判断材料として求めている項目を列挙することで、口コミにこれらの項目に該当する情報が1つでもあれば必要、そうでなければ不要と判定する。この項目を以下に示す。

- 訪問時間
- 店の営業時間
- 店の雰囲気・外観
- 来店客数(混み具合)・客層
- 料理が出てくるまでの時間
- 料理の見た目・味
- 店員の接客態度
- メニュー・価格
- 店舗の場所
- おすすめの品

結果として、集めた文のおよそ半分を必要な文として分けることができた。この2つの群に分けた口コミを用いて、必要な文、不要な文それぞれの特徴を明らかにする。

4.2 出現率を用いた評価値の算出

本研究では必要な文と不要な文の特徴は、文中の単語の出現頻度の違いによって現れると考えた。そこで、2つの群別に単語の出現頻度を求めた²。

2つの群に含まれる文の数は同じではないため、比較ができるように単語の出現数をそれぞれの群に含まれる文の総数で割った出現率を求める。つまり出現率とは、それぞれの群において1つの文あたりのその単語の出現割合を意味する。必要な文と不要な文における単語の出現率を求めた結果を表1に示す。

表1:要不要の群別の出現率上位の単語

必要な文に出現する単語			不要な文に出現する単語		
1位	パン	6.54%	1位	普通	2.52%
1位	美味しい	4.36%	1位	ラーメン	2.52%
3位	カレー	4.12%	1位	行く	2.52%
3位	クリーム	4.12%	4位	パン	2.29%
5位	店内	3.87%	5位	ランチ	2.06%
6位	スープ	3.63%	5位	今回	2.06%
7位	注文	3.15%	5位	カレー	2.06%
8位	普通	2.91%	8位	利用	1.38%
9位	ラーメン	2.66%	8位	注文	1.38%
9位	良い	2.66%	8位	ビール	1.38%
10位	雰囲気	2.42%	8位	思う	1.38%
11位	風味	2.18%	12位	帰り	1.15%
11位	コース	2.18%	12位	最初	1.15%
13位	盛り	1.94%	12位	以前	1.15%
13位	セット	1.94%	12位	以上	1.15%
13位	今回	1.94%	12位	会社	1.15%
13位	通り	1.94%	12位	予約	1.15%
17位	たっぷり	1.69%	12位	息子	1.15%
17位	過ぎ	1.69%	12位	素晴らしい	1.15%
17位	ソース	1.69%	20位	ごちそうさま	0.92%

表1は名詞、形容詞、動詞について出現率の高い上位20個の単語を降順で並べたものである。これらの差が小さな単語は群特有の単語ではないと考えられるため、2つの群の出現率の差が1%以上で、かつ出現率の上位の単語を特徴単語と定義する。表1において背景に色のついているものが特徴単語である。表2はこの特徴単語の出現率上位の一覧である。

次に、文単位で必要か不要かの判断をするための評価値の算出方法について述べる。評価値の考え方は、表1で求めた必要文の特徴単語の出現数が多いほど、その文は必要の可能性が高まり、逆に不要文の特徴単語が多く含まれるほど不要と判定される可能性

¹ <https://tabelog.com/>

² 形態素解析や頻度算出には、<https://textmining.userlocal.jp/> のサイトを利用した。

が高まるものとする。それゆえ評価値は、対象となる文中に現れる必要文の単語1つにつき+1、不要文の単語1つにつき-1を加えることで算出する。

また、特徴単語で一律に同じ値で加減算するのではなく、特徴単語の出現率に応じて重み付けした加重和によって評価値を算出する方法も提案する。これは出現率が高い特徴単語ほど重みを大きくする考え方である。具体的には、必要文の特徴単語として選択したものの中で、最も出現率が低い単語の重みを1とし、これに対する出現率の比で重み付けする。例えば、25個の特徴単語中で最も出現率が低い単語の出現率を0.5%とすると、出現率が1%の単語の重みは2となる。

さらに、評価値算出に用いる単語数による差異を検討するため、必要・不要の特徴語をそれぞれ15個(背景に色がついている単語のみ)用いた場合と25個(表中のすべての単語)を選び出した場合と2つの方法を評価検証する。

表2:特徴単語一覧

必要文の特徴単語				
パン	カレー	クリーム	店内	スープ
注文	雰囲気	風味	コース	盛り
セット	通り	過ぎ	美味しい	良い
たっぷり	ソース	満席	値段	個人的
お酒	ライス	購入	ボリューム	濃い

不要文の特徴単語				
ランチ	食事	利用	会社	帰り
最初	メニュー	以前	以上	予約
ごちそうさま	息子	行く	思う	素晴らしい
提供	有名	仕事	家事	ハンバーグ
オープン	近く	来る	飲む	多い

4.3 分析結果と考察

ロコミの文単位で要不要を判断する評価値の有効性を検証するために、まず30件の検証用ロコミデータを文単位で、人手によって必要な文と不要な文に分類した。これらの文に対して、人手で定めた要不要の判定結果と評価値に基づく判定結果が一致するか、その誤り率を検証した。評価値に基づく判定では、評価値が閾値を超えているか否かによって決定する。この閾値の設定値による変化についても検証した。

誤り率には2つのタイプがある。1つは、正解のデータでは必要としたが予測で不要だと判断されてしまった誤りであり、もう1つは本来不要のはずが必要だと判断されてしまった誤りである。前者の誤り率は、人手で判別した必要な文の総数を分母とし、必要な文を不要

と誤判定した文の数を分子として算出できる。後者の誤り率は、人手で判別した不要な文の総数を分母とし、不要な文を必要と誤判定した文の数を分子として求めることができる。

誤り率のグラフを図1から図4に示す。図1、図2は15個の単語を用いたときの結果であり、図3、図4は25個の単語を用いたときの結果である。また、図1、図3は重み付けを行っておらず、図2、図4が重み付けを用いたケースとなっている。

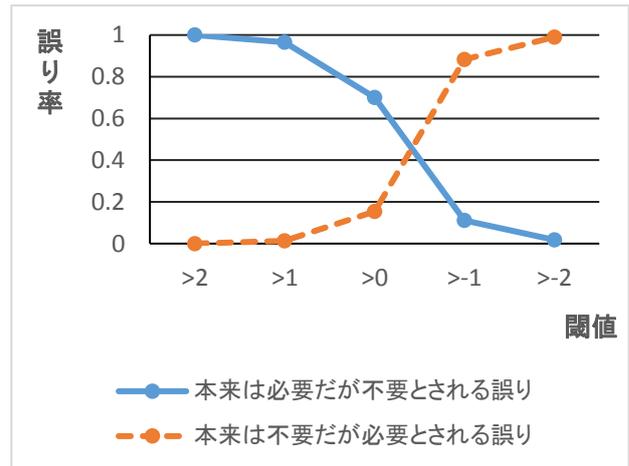


図1: 閾値による2つの誤りの割合の変化 (単語15個ずつ、重み付けなし)

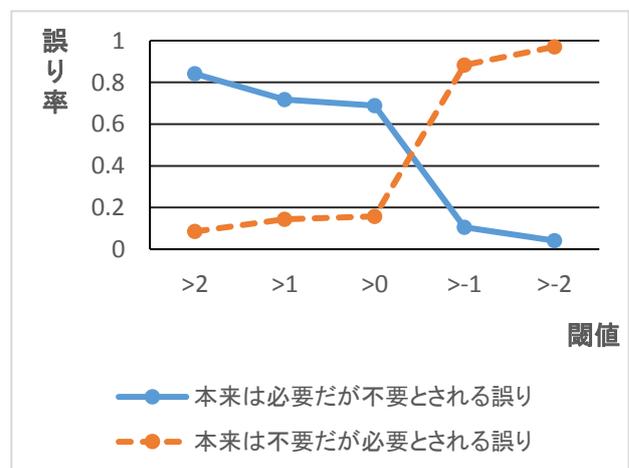


図2: 閾値による2つの誤りの割合の変化 (単語15個ずつ、重み付けあり)

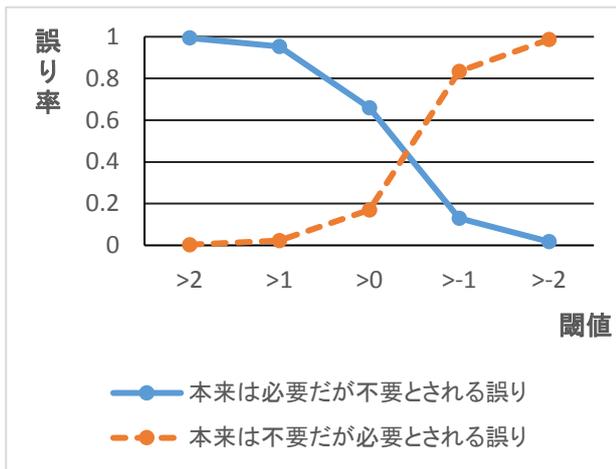


図3: 閾値による2つの誤りの割合の変化 (単語25個ずつ, 重み付けなし)

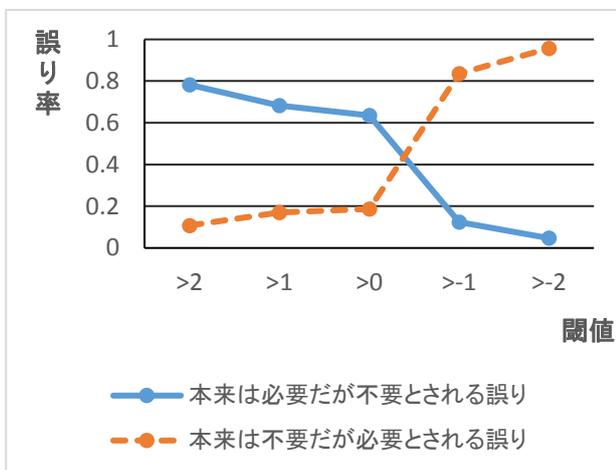


図4: 閾値による2つの誤りの割合の変化 (単語25個ずつ, 重み付けあり)

グラフから、閾値を小さい値にするほど本来必要であるものを不要と判断してしまった誤りの割合は低くなるが、本来不要な文を誤って必要と判断してしまった誤りの割合は逆に高くなることが読み取れる。また、2種類の誤り率の和で考えると、閾値を0とした場合が最も小さくなっていることがわかる。このことから、適切な閾値は0と考えられる。さらに、特徴単語15個ずつの場合と25個ずつの場合の比較では、25個の場合がわずかに全体的な値が小さくなっている。重み付けの有無の比較では、重み付けした評価値を用いた場合の誤り率が全体的に小さくなることがわかった。このことから、評価手法の1つとして文の有用性の判別に活用するためには、重み付けを行い、かつ単語の数を増やすことが最適であると考えた。

5. むすび

本研究では、食べログのロコミから単語の出現率を算出し、評価値に基づいて出現率の高い単語がロコミ文を判別するのに有用であることを2つの誤りの割合を算出することによって示した。その結果、グルメサイトのロコミにおいて出現率の高い単語は、文の有用性を判断するための評価指標になり得るということがわかった。

今後の課題として、本研究ではロコミを文ごとに分けて文を判別する際の単語の有用性を示すことを目的としたが、1つの文では含まれる単語、情報が少なく、出現率が高いとした単語がまったく現れない場合が多く見られた。このことから、文単位ではなく1つのロコミ単位で有用性を示す検証を行い、精度を高めることが重要であると考えた。そのために、文単位でロコミを必要、不要に判別する方法を検討、検証することが今後の課題として挙げられる。

参考文献

- [1] グルメサイト「4強時代」、生き残るのはどこか | 食品 | 東洋経済オンライン, 2016年11月24日
<http://toyokeizai.net/articles/-/144841?page=4>
- [2] グルメサイト・アプリに関する利用実態調査 - MMD研究所, 2015年7月30日
https://mmdlabo.jp/investigation/detail_1466.html
- [3] グルメサイト『食べログ』『Retty』『ぐるなび』『ホットペッパー』をアクセス分析的視点から比較してみた - 目黒に居る, ブランディングとリクルーティングのプロデューサーのブログ, 2016年4月14日
<http://hiryu.hatenablog.com/entry/2016/04/14/193236>
- [4] 広告(メーカー・団体様等向け)について[食べログ] user-help.tabelog.com/advertisement/
- [5] 毛中杰, 服部俊一, 高間康史, “価値観に着目したレビューの有用性評価手法の提案”, 日本知能情報ファジィ学会 ファジィ システム シンポジウム 講演論文集, 28, 0, pp.338-341, 2012
- [6] 桑田大徳, 岡誠, 吉村宏樹, 森博彦, “自由形式で記述されたレビューの評価視点別自動分類の提案”, 全国大会講演論文集, 2011, 1, pp.399-400, 2011