

レストランレビューにおける料理特有の評価表現の抽出と分析

鈴木 稜也 山西 良典 西原 陽子

立命館大学大学院情報理工学研究科

{is0313fv@ed, ryama@fc, nisihara@fc}.ritsume.ac.jp

概要 本稿では、レストランレビュー中の料理名周辺にある記述から料理特有の評価表現を抽出する手法を提案する。レストランレビューの活用が一般化してきていることで料理についてのウェブ上の評価情報は日々増加しているが、大量の評価情報を読み、ユーザの嗜好に合った料理を選択することは難しい。このような背景のもとで、シズルワードのような一般的な味評価の表現のみならず、ある料理に対して共起頻度が高い調理方法を表す単語や料理特有の評価表現を容易に取得可能にすることは、レビューの記述支援や理解支援につながると考えられる。係り受け解析だけでなく、料理名を含む文の周辺に存在する単語の重要度をハイブリッドに用いることで、料理の評価表現を獲得する。分析結果の考察から、提案手法によって抽出された料理特有の評価表現は、一般的な味表現に比べて料理の特性をよりよく捉えている可能性が高く、レビュー記述時の多様なニーズに応じた情報提供が可能になることが示唆された。

キーワード 食メディア情報処理, レビュー情報処理, 味の評価表現

1 はじめに

「☆5つ」や「お肉がおいしい」といったレストランレビューは、レストランを検索するユーザにとっても、レストランにとっても有益なレビューとならないことが多い。レストランを検索するユーザがレビューを閲覧する意図は、点数での評価や単純な感想ではなく、食べなくては分からない料理の詳細を推測するための具体的な説明を得ることもあって考えられる。そのため、具体性に乏しいレビューは、ユーザのレストラン選択の参考情報としての有用性は低い。昨今の飲食店にとってはレビューそのものがレストランの広告として機能している側面もある。すなわち、具体性の低いレビューは、レストランにとって広告効果が小さいとも捉えることができる。しかしながら、料理の様子を具体的に表現したレビューを書くためには、料理に対して適切な評価表現を用いる必要がある。Amano *et al.* [1] はインターネット上の健康管理および食事情報 Web サービスにおける頻出単語の分析から、専門家以外の一般の人にとってメニューについての情報を正しく記入することは困難であると述べている。

料理の評価には様々な評価表現が含まれており、特に、シズルワード¹と呼ばれる味表現が用いられることが多い。シズルワードは「おいしい」「食べたい」を感じるかを調査して定められた単語群であり、味覚系、食感系、情報系に細分化された計 362 語が用意されており、Web 上の様々なメディアで料理の評価に用いられている [2]。一方で、料理ごとの特有の評価表現も存在する。例えば、「ハンバーグ」のレビューでは、デミグラスソースや和

風ソースといったソースの種類や評価が、ハンバーグの具体的な味を想像させる上で役に立つ。料理の様子の説明や評価に有用な料理特有の表現が明らかになれば、レビューアに対して料理ごとにレビューすべき観点であったり、評価表現を示唆することが可能になると考えられる。

本稿では、レストランレビュー中の料理名周辺にある記述を分析することで料理特有の評価表現を抽出する手法を提案した。得られた分析結果についての考察から、レストランレビューの記述支援やレビューの理解支援への応用について検討する。

2 関連研究

Kato *et al.* [2] は、シズルワードを用いた料理の評価表現について、Twitter², Web ページ, ユーザ作成レシピサイト間で比較分析している。シズルワードはメディアによって使われている単語の種類が異なることを明らかにしており、Web ページでは情報系のシズルワード、Twitter では料理特有の評価表現がそれぞれ頻出することを示している。本稿では、レビュー情報からシズルワード以外のそれぞれの料理特有の評価表現を自動的に獲得するしくみを提案する。

Nobumoto *et al.* [3] は、海外での食事の際にメニューを読むことは可能であっても文化の違いから料理の詳細の推測が難しい点を問題として挙げている。観光客が旅行先での自身の理想の食事をするために、レストランメニューの多言語での表現を料理の特徴ベクトルを用いて実現している。本研究で獲得する料理特有の評価表現を異文化で比較することで、料理に対しての異文化での評価の差異が明らかになると期待される。

Copyright is held by the author(s).

The article has been published without reviewing.

¹<http://www.bmft.jp>

²<https://twitter.com>

表 1 分析対象の料理名一覧.

うどん	うなぎ	おでん
おはぎ	オムライス	お好み焼き
カツ丼	カフェラテ	カレー
きつねうどん	ケバブ	コーンポタージュ
クロquette	しゃぶしゃぶ	シューマイ
すき焼き	ステーキ	ぜんざい
チーズケーキ	チャーハン	テールスープ
とんかつ	ハヤシライス	ハンバーガー
ハンバーグ	バンバンジー	ピザ
ブイヤベース	ポテトサラダ	マルゲリータ
メンチカツ	もつ鍋	ラーメン
ローストビーフ	ワンタンスープ	串カツ
唐揚げ	天丼	寿司
春巻き	杏仁豆腐	焼き鳥
牛タン	牛丼	納豆
肉じゃが	親子丼	豚キムチ
餃子	麻婆豆腐	

3 分析方法

以下の手順に従ってレストラン検索サイトのレビューを分析する。表 1 に、分析対象とした料理の一覧を示す。

1. gurunaviAPI³を用いて、表 1 に示した料理名が含まれるメニューについてのレビューを上限 1,000 件として取得する。上限 1,000 件のレビューが取得できた料理名は表中下線文字であった。
2. 取得したレビューを文単位に分割する。その際、「,」「.」、記号で分割する。
3. 分析対象の料理名を含む文と、その前後 3 文の合計 7 文を取得する。
4. 手順 3. で取得したレビューを対象とし、料理名を示す単語に対し全文係り受け解析を行い、料理名を修飾している単語を得る。
5. 手順 3. で取得したレビューを単語単位に分割し、名詞、動詞、形容詞、形容動詞の 4 品詞に絞り込む。
6. 手順 5. で得られた単語に対して、TF-IDF 法を用いて単語の重要度を算出する。

手順 4. では、全文係り受け解析を自然言語処理ライブラリ「GiNZA」⁴を用いることで解析処理をする。解析処

理を行った結果、形態素解析された単語とそれぞれ修飾している単語が明らかになる。手順 5. では単語単位に分割する際、形態素解析は MeCab を、辞書は NeologD を使用する。手順 6. では、TF-IDF により算出された重要度を用いて、他の料理のレビューにはあまり出てこないが、注目している料理のレビューにのみ頻出する単語を抽出する。重要度指標 $tfidf(t, d)$ は下式で算出される。

$$tfidf(t, d) = \left(\frac{f_{t,d}}{\sum_{s \in d} f_{s,d}} \right) \times \left(\log \frac{N}{df(t)} + 1 \right), \quad (1)$$

ここで、 $f_{t,d}$ は手順 5. で得られた単語 t の任意の料理のレビュー d 中での頻度を示し、 $df(t)$ は単語 t がレビュー中で出現する料理数を示す。 N は分析対象とする全料理数 (本稿では $N=50$) を示す。

本手順に従うことで、分析手順 4. から「料理を修飾している単語」、分析手順 6. から「料理名周辺で特徴的に現れる単語」がそれぞれ獲得される。本稿では、これらの単語の使われ方と料理の評価としての妥当性について考察する。

4 分析結果

本稿では、表 1 中で下線文字で示した料理、つまり、「うどん」「カレー」「ステーキ」「ハンバーグ」「ラーメン」「餃子」を分析対象とする。表 2 に、分析手順 4. によって獲得された料理名を修飾している単語について、分析対象とした料理名毎の頻出上位 5 単語を示す。表 2 の結果から、一般的な味表現とされるシズルワードはほとんど存在しないことが明らかになった。例えば、ハンバーグのレビューで多く見られた表現としては「デミグラスソース」が頻出単語として抽出された。デミグラスソース自体は味表現ではなく、ハンバーグの属性を示す単語であり、料理の詳細を評価するための単語であると示唆される。ハンバーグはソースの種類によって味が大きく変わるため、ソースの種類は重要な評価表現として捉えられる。餃子のレビューでは調理方法である「水」「焼き」「蒸し」が料理名に関連することが多かった。調理方法も料理ごとに大きく変化するものであり、餃子の味や食感を推測するための評価表現として重要であると考えられる。料理名を修飾する単語について、全体的な傾向として評価に用いられることが多いことと、調理方法や味付けの表現が多いことが明らかになった。

表 3～表 10 に、分析手順 6. で得られた料理名周辺で特徴的に現れるシズルワードとシズルワード以外の単語を 10 単語ずつ示す。表中、単語の頻出順の昇順で示しており、最も高い重要度を示した単語については太字で示した。

表 3 から、「ラーメン」についてのシズルワードの中で

³<https://api.gnavi.co.jp/api/>

⁴<https://megagonlabs.github.io/ginza/>

表2 料理名毎の係り受け解析によって得られた料理名を修飾している単語の頻度上位5件.

	うどん	カレー	ステーキ	ハンバーグ	ラーメン	餃子
1位	カレー	チキン	ソース	ジューシー	味噌	水
2位	讃岐	インド	サラダ	煮込み	塩	焼き
3位	かけ	カツ	ボリューム	チーズ	とんこつ	ジューシー
4位	煮込み	グリーン	スープ	デミグラスソース	博多	蒸し
5位	冷たい	スープ	柔らかく	ソース	豚骨	皮

表3 ラーメンに関する TF-IDF 値 (シズルワード上位10件).

評価表現	TF-IDF 値
ボリューム	0.16×10^{-2}
濃厚	0.70×10^{-2}
シンプル	0.20×10^{-2}
絶妙	0.09×10^{-2}
ジューシー	0.02×10^{-2}
新鮮	0.04×10^{-2}
あつあつ	0.00×10^{-2}
ヘルシー	0.05×10^{-2}
まろやか	0.20×10^{-2}
贅沢	0.04×10^{-2}

表4 ラーメンに関する TF-IDF 値 (シズルワード以外の単語上位5件).

評価表現	TF-IDF 値
スープ	1.73×10^{-2}
麺	2.07×10^{-2}
チャーシュー	1.34×10^{-2}
醤油	0.64×10^{-2}
味噌	0.48×10^{-2}

最も高い重要度 0.70×10^{-2} を示した単語は「濃厚」であった。また、シズルワード以外の単語については、表4中で最も高い重要度 2.07×10^{-2} を示した単語は「麺」であった。表5と表6から、「餃子」についての最も高い重要度 1.07×10^{-2} を示したシズルワードは「ジューシー」であり、最も高い重要度 1.66×10^{-2} を示したシズルワード以外の単語は「皮」であった。シズルワードとシズルワード以外の単語について最も高い重要度を比較してみると、シズルワード以外の単語がシズルワードに比べて「ラーメン」で約3倍、「餃子」で約1.5倍高い値を示した。表6では、「ビール」が高い値を示している。「ビール」は「餃子」の材料でもなく、調味料でも調理方法でもない。本研究で実際に使用したレビューに、「中国出張の帰りにふらっと立ち寄りました。カウ

表5 餃子に関する TF-IDF 値 (シズルワード上位10件).

評価表現	TF-IDF 値
ボリューム	0.31×10^{-2}
濃厚	0.02×10^{-2}
シンプル	0.26×10^{-2}
絶妙	0.19×10^{-2}
ジューシー	1.07×10^{-2}
新鮮	0.02×10^{-2}
あつあつ	0.26×10^{-2}
ヘルシー	0.14×10^{-2}
まろやか	0.00×10^{-2}
贅沢	0.04×10^{-2}

表6 餃子に関する TF-IDF 値 (シズルワード以外の単語上位5件).

評価表現	TF-IDF 値
皮	1.66×10^{-2}
水餃子(水)	1.31×10^{-2}
肉汁	0.54×10^{-2}
焼き餃子(焼き)	0.76×10^{-2}
ビール	0.30×10^{-2}

ンターのみでこじんまりとした感じの店です。まずは餃子二枚とビールと一品のセットでスタートです！」というレビューが存在した。このレビューからもわかるように、中華料理レストランではビールと餃子をセットで販売している店舗も少なくはなく、これらの料理を同時に食事することが推察される。任意の料理のレビュー中に他の料理名が出現し、その料理名の重要度が高ければ代表的なサイドメニューとして捉えられる可能性が示唆された。

「肉料理」という同一ジャンルという観点で、各料理についての分析結果を比較した。表7と8に示したステーキの周辺単語と表9と表10のハンバーグの周辺単語について考察すると、「ソース」に関わる単語の重要度が高い傾向が見られた。一方で、ラーメンやうどんでは「ソース」についての重要度は高くなく、料理のジャンル

表7 ステーキに関する TF-IDF 値 (シズルワード上位 10 件) .

評価表現	TF-IDF 値
ボリューム	0.51×10^{-2}
濃厚	0.20×10^{-2}
シンプル	0.14×10^{-2}
絶妙	0.16×10^{-2}
ジューシー	0.40×10^{-2}
新鮮	0.20×10^{-2}
あつあつ	0.07×10^{-2}
ヘルシー	0.10×10^{-2}
まるやか	0.04×10^{-2}
贅沢	0.33×10^{-2}

表9 ハンバーグに関する TF-IDF 値 (シズルワード上位 10 件) .

評価表現	TF-IDF 値
ボリューム	0.54×10^{-2}
濃厚	0.53×10^{-2}
シンプル	0.10×10^{-2}
絶妙	0.18×10^{-2}
ジューシー	0.84×10^{-2}
新鮮	0.06×10^{-2}
あつあつ	0.17×10^{-2}
ヘルシー	0.23×10^{-2}
まるやか	0.06×10^{-2}
贅沢	0.15×10^{-2}

表8 ステーキに関する TF-IDF 値 (シズルワード以外の単語上位 5 件) .

評価表現	TF-IDF 値
ソース	0.65×10^{-2}
サラダ	0.54×10^{-2}
リーズナブル	0.36×10^{-2}
鉄板	0.43×10^{-2}
焼き	0.38×10^{-2}

表10 ハンバーグに関する TF-IDF 値 (シズルワード以外の単語上位 5 件) .

評価表現	TF-IDF 値
ソース	1.17×10^{-2}
肉汁	0.92×10^{-2}
チーズ	0.68×10^{-2}
デミグラスソース	0.86×10^{-2}
レア	0.40×10^{-2}

ごとに重要となる単語が共通する可能性が示唆された。

5 おわりに

本稿では、メニューの詳細を推定可能な具体性をもったレビューの記述支援を目標として、レストラン検索サイトのレビューを分析することで料理ごとの評価表現を抽出する手法を提案した。料理ごとに特徴的に出現する単語はシズルワードと並んで重要な単語と考えられる。また、料理ごとに特徴的に出現する単語には調理方法や調味料などの料理の詳細と一緒に食事することによって味を引き立てるサイドメニューなどが多いことが示唆された。得られた単語をレビュアーにサジェストするしくみを構築すれば、メニューの詳細を示した具体性の高いレビューの記述支援につながると考える。

今後は、分析対象の料理数を増やすと共に、他言語レビューでの分析も検討する。また、レビュー記述支援システムを構築する。

謝辞

本稿では、ぐるなび Web サービスから提供されたレビュー情報を用いた。本研究は、一部、すかいらくフードサイエンス研究所の研究助成のもと行われた。記して、謝意を表す。

参考文献

- [1] Sosuke Amano, Kiyoharu Aizawa, and Makoto Ogawa. Frequency statistics of words used in japanese food records of foodlog. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 2014.
- [2] Daisuke Kato, Mai Miyabe, Eiji Aramaki, and Akiyo Nadamoto. Sizzle word analysis from multiple internet media. In *2015 18th International Conference on Network-Based Information Systems*, 2015.
- [3] Kensuke Nobumoto, Daiju Kato, Masaki Endo, Masaharu Hirota, and Hiroshi Ishikawa. Multilingualization of restaurant menu by analogical description. In *Proceedings of the 9th Workshop on Multimedia for Cooking and Eating Activities in Conjunction with The 2017 International Joint Conference on Artificial Intelligence*, CEA2017, pp. 13–18, 2017.