

口コミ解析における識別精度の向上方法の提案と 解析システムの構築

A proposal of accuracy improvement method and a construction of the analysis system in review analysis

上谷竜士[†], 菱田隆彰^{††}
Ryuji UETANI[†], Takaaki HISHIDA^{††}

Abstract In recent years, comments posted on review sites have a strong influence on viewers' willingness to purchase. In order not to be misled by wrong information, it is necessary to read more comments on favorite items. However, it takes a lot of time to read all posted comments. In this paper, we focus on emotional words included in the review, build a system to analyze the comments of the review sites, and propose a method to improve comment classification accuracy for the analysis system.

1. はじめに

近年口コミサイトに投稿されるコメントが閲覧者に強い購買意欲を与えるようになった。閲覧者は興味のある対象物の口コミから「満足」や「好き」などといった感情を示す単語（以降感情語と示す）に着目して対象物の雰囲気把握している。しかし、対象物の正確な感情傾向を把握するには全ての口コミを精査した上で客観的な分析を行う必要があり、時間と手間が必要になる。

我々は杉本ら[1]によって提案された人の感情を示す単語を収録した辞書（以降杉本式辞書と示す）を用いて価格.com, 食べログ, TripAdvisorに投稿される口コミを解析し、解析結果を可視化する口コミ解析システムの試作を行った[2]。試作した解析システムを利用する手段としてスマートフォンアプリを製作した。

試作した解析システムを利用することで利用者は気軽に対象物の印象を把握することが可能となった。しかし、口コミに含まれる一部の感情語は文章の内容によって感情以外の意味を示す単語に変化することがあり、その識別方法が課題であった。

本研究では、口コミ解析に適した感情語辞書と教師データの検討を行い、感情語の意味を識別する機能として機械学習による実装と精度検証を行う。

また、解析結果の表示方法を再検討し、スマートフォンアプリ以外の利用方法としてパソコンの Web ブラウザから利用できる WebAPI を用いた解析結果表示機能の実装を行う。

2. 関連研究

Web上に投稿される口コミやブログ記事などの文章が示す評判や雰囲気を判別する研究は目的、手法も様々に行われている。鈴木ら[3]の研究では、Web上から収集した評判情報が肯定的であるか、否定的または非評価的な評判であるか判定を行う際に機械学習モデルであるナイーブベイズ分類器と EM アルゴリズム, Support Vector Machine を組み合わせた独自のアルゴリズムを実装することで評判情報の分類を行っている。ナイーブベイズ分類器を中心として3種類の機械学習モデルを組み合わせることで正答率が向上することを示している。また、教師データ量を増やすことで分類正答率が上昇することを検証しており、十分な教師データ量を用意することが重要であることを示している。また、小林ら[4]の研究では、ネットショッピングサイトに投稿される口コミを収集し閲覧することができる Web アプリケーションの開発・提供を行っている。収集した口コミがポジティブまたはネガティブな文章であるか分類する際にナイーブベイズを用いて分析機能の実装と実証実験を行っている。実験結

[†] 愛知工業大学大学院 経営情報科学研究科 (豊田市)

^{††} 愛知工業大学 情報科学部 (豊田市)

果では分類正答率が 50%前後を示したことから, 約 5 割のロコミを正しく分類していることがわかる。

3. ロコミ解析の手順と感情語辞書

3・1 ロコミ解析の手順

ロコミサイトに投稿されたコメントには, 投稿者の主観的な意見が述べられており, 多くの場合感情的な単語が含まれる文章によって表現されている。投稿閲覧者は数多くの投稿者の感情的な意見を解釈し, 自身の購買行動を決めるための情報の一つとして利用する。コメントには判断を誤らせる不正な情報が含まれる場合があり, 判断ミスを減らすためには, 多くのコメントをよく読み様々な意見を咀嚼する必要がある。

しかし, 人気のある対象物には, 多くの投稿が寄せられており, 投稿されたコメントを全て読むには長い時間を要する。ノイズの少ない必要な要点が示される情報に精錬するには, 全てのコメントから適切に要約された分かりやすい値を算出する手順が求められる。そこで我々は, コメント内に含まれる感情語に注目する。どのような感情語がどの程度含まれるかを数値化することで, その対象物が投稿者全体からどのような印象を持たれているかを端的に示すことができると考えた。

表 3.1 杉本式辞書の収録例

カテゴリ	収録数	収録例
喜び	268	嬉しい, 満足, 感謝
楽しさ	106	楽しい, 興じる, 面白い
安心	118	落ち着く, 気楽, 和める
好き	299	大好き, 愛情, 慕う
驚き	152	驚愕, 衝撃, 狼狽

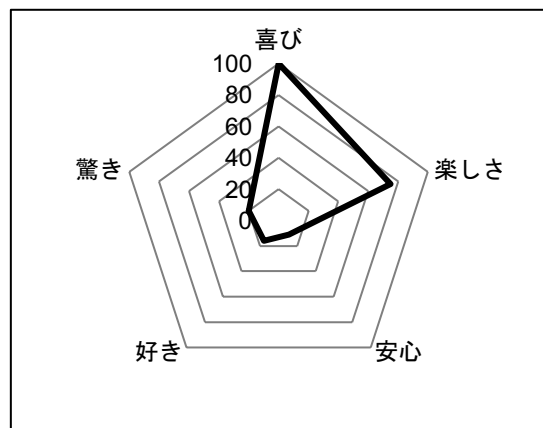


図 3.1 解析結果の可視化

先行研究として杉本ら[1]は, ロコミサイトに投稿されているコメントの調査を行った。また, 感情表現辞典[5]に収録されている感情語および感情語の分類を元に, ロコミの分類に適した 5 つの感情語のカテゴリ「喜び」「楽しさ」「安心」「好き」「驚き」を提案し, 各カテゴリに含まれる感情語を振り分けた感情語辞書(以降, 杉本式辞書と呼ぶ)を作成した。表 3.1 に杉本式辞書の収録例を示す。辞書の内容は感情表現辞典に掲載されている単語を整理し, カテゴリのラベル付けを行ったものである。

我々は文献[2]において, この杉本式辞書を用い, ロコミ解析システムの試作を行った。その解析手順を以下に示す。

1. 対象物に投稿されたコメントを入手する
2. コメント文を形態素に分割する
3. 形態素に分けた単語の中で感情語辞書に含まれる語を抽出する
4. 抽出した感情語を辞書に示されたカテゴリ毎に出現数を集計する
5. 集計された 5 つの値の中の最大値を 100 として正規化を行う
6. 正規化された 5 つの値を対象物の印象を表す特徴量とする

例えば, ある商品に投稿されたコメントについて, 「喜び」「楽しさ」「安心」「好き」「驚き」の順に各カテゴリに含まれる感情語の出現数が [60, 45, 7, 10, 12] であった時, その商品の印象を表す特徴量は [100, 75, 11, 16, 20] となる。この特徴量を使用し, 図 3.1 のようなレーダーチャートで表すことで, ロコミ全体が示す対象物の印象を閲覧者は一目で把握することが可能となる。

3・2 修正辞書の作成

ロコミ解析の際の感情語の抽出には杉本式辞書を使用する予定であったが, 使用するうちにいくつかの問題点が生じた。一つは, 収録単語が漢字で表現されたものを中心に収録されている点である。ロコミサイトに投稿されるロコミは漢字で表現できる単語をひらがなやカタカナで表現するケースが多く, 調査したところ抽出の際に感情語の取りこぼしが多量に発生していることが分かった。二つ目は, 収録単語の一部は辞書で意図された内容とは異なる意味で広く使用されている場合が存在する点である。例えば「楽」という単語は, 辞書では楽しさを表す意味として分類されているが, ロコミ内では「こんなに楽なんだ」「使えば楽ですね」のような使い方が多く, 「楽しい」に関係する感情より「安心」を示す感情語として出現するケースが多く見受けられた。

そこで我々は, 杉本式辞書を元に単語の追加・変更を行った修正辞書を作成した。主な変更点を以下に示す。

1. 収録されている感情語のひらがな, カタカナ表現(以降, かな表現と表す)の追加
2. 既存の単語の分類先の変更

1のかな表現の追加については, 一部の表現はロコミ内で他の意味として用いられることがあり, その利用状況を確認し, 登録を行う必要がある. 代表的なかな表現の一部を表3.2に示す. 「どう」に関しては最も出現数が多く主に「どうでしょうか」, 「どうしたらいいかわからない」など副詞となる「どう」が抽出される場合が目立った. 「アンド」では“&”をカタカナで表現される場合に用いられていた. 「かんじ」や「とうぜん」も「感じ」「当然」のひらがな表現として利用されていた. いずれの場合も, 本来の意味としては使用されることはほぼ無く, 辞書への追加は行わないこととした.

作成した修正辞書の収録例および杉本式辞書と修正辞書を用いたロコミ内の単語抽出数の比較例を表3.3, 3.4に示す. 収録語数が増えたことで抽出量が格段に増加したことが分かる. 追加した単語は, ロコミの特性に合わせた調整がされており, 適切な単語をより多く抽出可能となった. これまでよりも精度の高い解析が期待できる. 以降は修正辞書を用いてシステムの構築, 評価を行う.

表 3.2 誤判定されるかな表現の例

元の感情語	誤判定されるかな表現
動	どう
安堵	アンド
莞爾	かんじ
陶然	とうぜん

表 3.3 修正辞書の収録例

カテゴリ	収録数	収録例
喜び	490	幸せ, こちよい, ウレシイ
楽しさ	219	鼓動, うずうず, オモシロイ
安心	221	楽, くつろぐ, ノンキ
好き	573	うっとり, 好み, ユウジョウ
驚き	195	意外, あんがい, ヨソウガイ

表 3.4 2つの辞書による抽出単語数の違い

カテゴリ	杉本式辞書	修正辞書
喜び	239730	464741
楽しさ	86293	112134
安心	34359	123102
好き	167177	262814
驚き	37921	108449

4. 識別機能の検討

4.1 識別機能の追加

3.1節に示した解析手順は, ロコミの内容に含まれる単語の内, 用意した辞書に含まれる単語に適合したものを抽出し集計する, いわゆるパターンマッチングによる抽出手法である. この方法では, 内容によっては意味のことなる単語であっても適合した単語として集計されてしまい, 誤差として解析結果に影響を与えてしまう.

我々が調査したところ, ロコミの内容は複雑な文法や比喩表現が頻繁に使用されることはないことが分かっており, 否定的な言い回しが出現する割合も小さい. そのため, パターンマッチングによる単語抽出だけでも解析結果が大きく異なる傾向を示すことは少ないと考えられる. しかし, 一部の対象物はコメント数が少なく, 一つの文章が全体の印象に大きく影響を与える場合があり, 対象物以外の話題や否定的な意見など本来抽出対象にならない文章をできる限り減らすことが望ましい.

そこで我々は, 解析手順の過程において, 文章を形態素に分解した後, その文章が集計するべき対象かどうかを判別するための識別機能を追加することを検討する. 今回は機械学習による識別機能の実装を行う. 機械学習モデルには, 文書分類によく利用され効果の高いとされるナイーブベイズ分類器を用いる.

4.2 教師データ作成方法の提案

ナイーブベイズによる分類には教師データが必要となる. 一般的にロコミ文に含まれる単語全体の集合を分類するカテゴリに対応付けして教師データとして用いることが多いが, 本研究では, 感情語とその周辺の内容との関係性が重要であるため, 文章中の幾つかの単語に絞って教師データの構成を検討する.

本節では, 二つの教師データの構成について提案を行う. 一つ目は, 文章中の感情語と係り先の単語の2つの単語を組として一つの教師データとする方法である. 以降, この教師データの構成を2組パターンと呼ぶ. 2組パターンによる教師データの作成例を図4.1に示す. 「画面が驚くほど綺麗に見えます」という文では, 感情語である“驚く”が“綺麗”に係っており, ロコミ投稿者は画面に対して想像以上に綺麗に見えることに驚いている

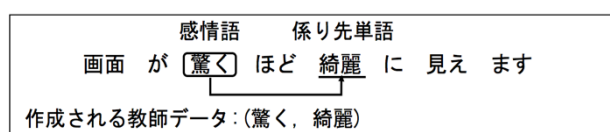


図 4.1 2組パターンの作成例

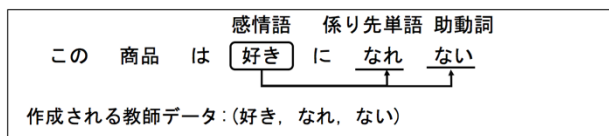


図 4.2 3 組パターンの作成例

ことがわかる。感情語に係っている単語を対象物の主要な要素かどうかで、分析対象とするかを識別できることが期待できる。

もう一つの方法は、前述の 2 組パターンの単語の組に 3 つ目の単語として感情語の後に出現する形容詞または助動詞を加えて教師データとする方法である。この構成を 3 組パターンと呼ぶ。3 組パターンによる教師データの作成例を図 4.2 に示す。「この商品は好きになれない」というロコミ文では感情語である“好き”という単語の係り先は“なれ”である。“なれ”という単語は「好きになれる」または「好きになれない」と表現することができ、2 組パターンでは情報が少なく識別が困難になる。しかし、3 組パターンを使用することで“好き”のあとに出現する助動詞の“ない”を情報として加えることで否定的なロコミ文と正しく判断することが可能になる。

4・3 識別機能の検証方法

提案する教師データの識別精度を検証する。検証に用いるデータは 2017 年 1 月から 2017 年 3 月の間に価格.com にロコミが投稿された対象物について、無作為に 300 商品選出し未知データとして用いる。未知データは予め手作業でカテゴリをラベル付しておく。また、教師データの作成にはロコミが 100 件以上登録されている対象物 100 商品が無作為に選出し、約 4000 個の教師データ群をそれぞれ作成する。教師データは提案する 2 組パターン、3 組パターン以外に、文章内の助詞を除いた単語群で教師データを構成したものを比較対象として用意する。以降、この教師データの構成をロコミパターンと呼ぶ。3 つの教師データに対して以下の手順で識別の正答率を求める。

1. 予め用意した 300 商品の未知データから無作為に 100 商品選択し未知データとする
2. 用意した教師データの内 3000 を無作為に選び学習させ未知データの識別を行う
3. 分類器によって算出された確率（以降、識別率と示す）が一定のしきい値（以降、識別基準値と示す）を超えたとき、その文章に含まれる感情語が対象物の印象を正しく示していると判断する
4. 対象物の印象を示す感情語であると正しく識別し

たときは TP、誤って識別したときは FP、対象物に対して否定または感情以外の意味を示す単語であると正しく識別したときは TN、誤って識別したときは FN として集計を行う

5. 以下の式から 1 商品毎の正答率 S を算出し 100 商品分行う

$$S = \frac{TP + TN}{TP + FP + FN + TN} \times 100$$

1 から 5 までの手順を試行回数 1 回として正答率を求め、それを 50 回実施しその分布を比較する。識別基準値については、10% から 90% まで 10% 刻みでそれぞれ分布の算出を行う。

4・4 検証と考察

3 つの教師データの構成、2 組パターン、3 組パターン、ロコミパターンのそれぞれにおいて、識別基準値毎の正答率分布を箱ひげ図としてそれぞれ図 4.3~4.5 に示す。横軸はそれぞれ設定した識別基準値、縦軸は試行 1 回あたりの正答率を示し各グラフはその識別基準値における 50 回分の試行の正答率の分布を表している。

今回使用した未知データは、感情語として正しく分類されるべき文章の割合は 79.17% であった。既存のパターンマッチングによる解析を行った場合、約 2 割は不正な文章を抽出してしまうことになる。

2 組パターンを用いた場合、識別基準値が 50% から低下し始め、以降急激に正答率が低下している。このことから正答とすべき感情語の算出される識別率が 40% 付近から分布し始めており、50% から 80% の値を示すものが一定量存在していることがわかる。今回調査した識別基準値としては 30% に設定することで試行 50 回の平均正答率が最も高く 88.99% となった。

3 組パターンを用いた場合、識別基準値が 60% から低下が始まり、70% を超えたあたりで急激に正答率が低下した。2 組パターンと比較して、正答率の落ち込みがより高い基準値から始まっており、正答とすべき文章がより高い値として算出されるようになっていることが分かる。教師データに追加した 3 つ目の情報が良く機能していると考えられる。平均正答率の最高値は 90.27% となり、その時の識別基準値は 2 組パターン同様 30% と設定した場合であった。

ロコミパターンを用いた場合、識別基準値 10% から正答率がなだらかに低下しており、正答の分布がかなりばらついていることが分かった。識別基準値を 10% に設定しても平均正答率は 68.91% であり、従来手法より値が低くなった。文章全体を教師データとした場合、教師データの基準が複雑になりすぎ、却って判断が曖昧になってしまったと考えられる。今回のような状況では、教師データの内容を適切に整理することで高い精度の識別が可能であることが示された。

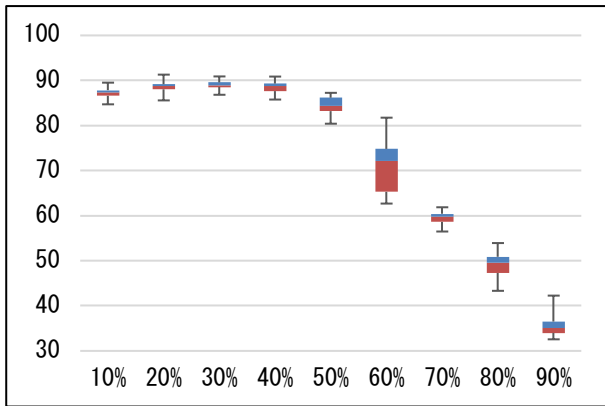


図 4.3 2組パターンの正答率の分布

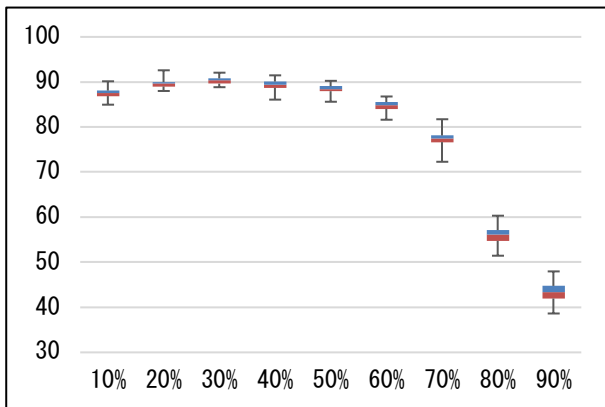


図 4.4 3組パターンの正答率の分布

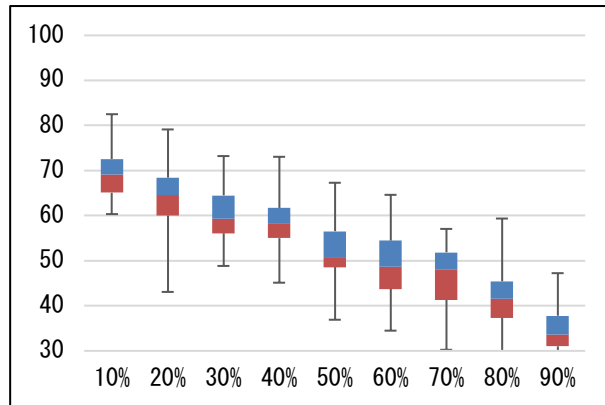


図 4.5 口コミパターンの正答率の分布

表 4.1 3組パターンで正しく識別される例

識別内容	感情語	否定・非感情語
口コミ	真鍮を使用した質感が好きで購入しました。	欲を言えば、注文住宅の様に好きな設定させてよって感じですかね
識別対象	好き, 購入, で	好き, 設定, な
識別率	54.31%	12.59%

3組パターンを用いて識別基準値を30%に設定したときに本識別機構が正しく機能する例を表4.1に示す。例にあげた口コミ文にはどちらも“好き”という感情語が含まれている。“好き”は口コミにおける出現頻度が多い感情語である。しかし、文章の内容によって様々な意味に変化しやすい感情語であり、従来の手法では意味の違いを識別することができず解析精度が低下する要因の1つとなる。

表の左側の文章は感情語として解析対象とすべき文章であり、右側の文章は対象物に対する“好き”ではないため対象から除外したい文章である。3組パターンの識別において、それぞれの文章ではそれぞれ[好き, 購入, で], [好き, 設定, な]の単語の組によって識別率が計算される。算出結果は54.31%と12.59%となり明確な差として表され、正しく仕分けられることになる。

5. 可視化機能の実装

5.1 機能概要

本章では、提案した解析手順を実現する口コミ解析システムの実装を行う。解析結果の可視化する機能は、より汎用的に利用することができるよう解析要求はWebAPIとして受け付け、結果をWebページとして返すよう実装を行う。また、利用者がWebブラウザから簡便に利用できるように、ブラウザの機能拡張モジュールによるインターフェースも用意する。今回は、Chromeブラウザを対象として、機能拡張モジュールの作成を行う。

実装する可視化機能の画面遷移図を図5.1に示す。画面は、「対象物一覧画面」、「解析結果画面」の2種類で構成する。解析可能なWebページを表示した時、解析用の拡張機能が有効となる。利用者が興味のある対象物の解析を行いたい場合、ブラウザ上部に配置されている拡張

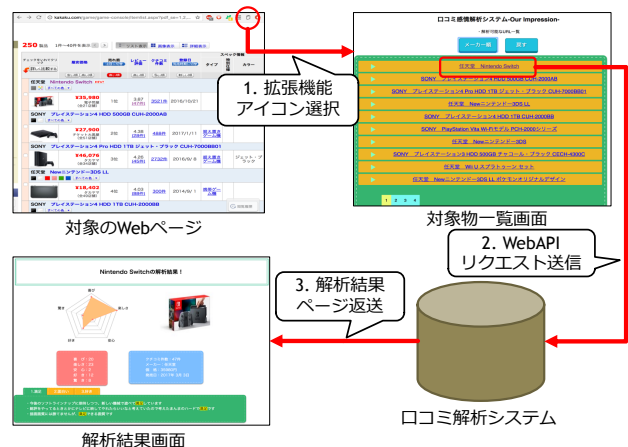


図 5.1 画面遷移図



図 5.2 対象物一覧画面

機能アイコンを選択することで対象物一覧画面がポップアップウィンドウとして表示される。対象物一覧画面から対象物を選択することで本口コミ解析システムに WebAPI リクエストとして解析要求が送信される。解析が終了すると解析結果は Web ページとして整形され、ブラウザに返送される。ブラウザは別のタブで解析結果画面が表示する。解析結果はそれぞれ別のタブで表示されるため、利用者は簡単に解析結果を比較できる。

5・2 対象物一覧画面

対象物一覧画面は、ブラウザの機能拡張アイコンが押された時に、表示されている Web ページから解析することが可能な対象物を収集し、一覧として表示する。利用者は対象物を選択することで口コミ解析を要求し結果を得ることができる。

図 5.2 では価格.com のゲーム機ランキング一覧ページに対して、商品の名称、メーカー名、各商品の口コミが掲載される URL を取得し表示した例である。取得した商品のリストを元に解析可能な対象物の一覧を 10 件ごとにページを分けて表示する。10 件を超える場合は画面左下にあるタブによって一覧の表示の切り替えを行うことができる。また、いくつかの項目に対する整列機能を実装しており、図の例ではボタンを選択することで対象物の表示順をメーカー毎に並び替えることができる。

5・3 解析結果画面

解析結果画面は、口コミ解析システムから生成される解析結果の Web ページであり、解析要求のあったブラウザに返され、別ページとして表示される。

図 5.3 は解析結果画面の一例である。画面は主に 3 つの要素で構成される。1 つ目は画面左上のレーダーチャートであり、前述した口コミ解析手法によって解析され

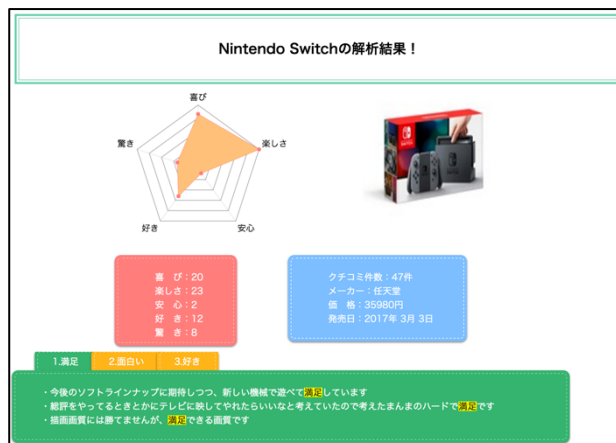


図 5.3 解析結果画面

た各感情カテゴリの特徴量を可視化したグラフである。5 つの感情のカテゴリ「喜び」、「楽しさ」、「安心」、「好き」、「驚き」に分けられた感情語の特徴量を示している。また、グラフの下には各カテゴリの感情語の出現数を併せて表示している。2 つ目は画面右上の対象物に関する情報である。解析可能なサイトである価格.com では対象物のイメージ画像、口コミ件数、メーカー名、最も安く購入できる価格を抽出し、表示する。3 つ目は、画面下部の解析中に得られた特徴的な口コミの一覧である。この一覧には口コミ解析中に多く出現した感情語を上位 3 つまで選出し、それぞれ感情語が含まれる口コミ文を 3 つまで無作為に選出し表示する。表示する文章は含まれる感情語に黄色のハイライトを付加し、利用者がその位置をひと目で把握できる。

本結果画面によって、対象物の基本的な情報と解析結果による口コミ全体が示す印象、その印象に強く影響を与える個々のコメントを一画面で確認することが可能となる。利用者が口コミ全体の意見を参考としながら、購入の検討や商品の比較が簡単に行うことができる。

6. まとめ

本研究では、口コミサイトの利用者が商品などの購入の際に利用する投稿されたコメントの読み取りと内容を把握・整理する作業の効率化を目的とし、投稿されたコメントの解析手法を提案し、解析の際に必要な口コミサイトに適した感情語辞書の作成、より高い精度で特徴量を抽出するための識別機構の検討を行った。また、構築した口コミ解析システムを汎用的に利用できる手段として WebAPI を用いた可視化機能の実装を行った。

解析に用いる辞書は修正を施すことで、既存の辞書に比べより多くの単語を適切に収集することが可能となっ

た。解析対象を判別するための機械学習による識別機構を検討し、従来より約 10%高い精度で識別可能な教師データの構成方法の一つを提案した。可視化機能ではブラウザの拡張機能から解析システムを利用可能とし、解析結果画面で検討に必要な情報が全て一画面内に表示されることで、利用者が簡単に比較検討が可能な環境を実現した。

今後の課題としては、商品とは異なる傾向を持つ観光地や飲食店の口コミサイトでの識別機能の精度検証を行いたいと考えている。

参考文献

- [1] 杉本祐介, 佐藤太一, 土井千章, 中川智尋, 太田賢, 稲村浩, 内藤克浩, 水野忠則, 菱田隆彰: 口コミを利用したレコメンドに適した感情語の分類方法の検討, ユビキタスコンピューティングシステム(UBI), 2015-UBI-45(50), pp. 1-6, 2015.
- [2] 松井瑠偉人, 上谷竜士, 梶克彦, 内藤克浩, 水野忠則, 菱田隆彰: 感情を活用した口コミ解析システムの実装, 情報学ワークショップ WiNF2015 講演論文集, pp.10-13, 2015.
- [3] 鈴木泰裕, 高村大也, 奥村学: Semi-Supervised な学習手法による評価表現分類, 言語処理学会第 11 回年次大会, 2005.
- [4] 小林亮, 鈴木浩, 服部哲, 速水治夫: ポジティブ・ネガティブ分析を備えたユーザレビュー集約システムの提案, 研究報告グループウェアとネットワークサービス, 2013-GN-87(1), pp.1-8, 2013.
- [5] 中村明: 感情表現辞典, 東京堂出版, 1993.

(受理 平成 30 年 3 月 10 日)