

Title	Twitterデータを用いた商品価値モデルによるレビュー 検索手法
Author(s)	鄂, 香伶
Citation	
Issue Date	2021-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/17088
Rights	
Description	Supervisor: 由井 蘭 隆也, 先端科学技術研究科, 修士 (情報科学)

修士論文

Twitter データを用いた商品価値モデルによるレビュー検索手法

E, Xiangling

主指導教員 由井 蘭 隆也准教授

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和 3 年 3 月

Abstract

Recent years, with the popularization of online shopping, there are a large amount of data on product evaluation on the Internet. Because these product reviews are expressed from the consumer's standpoint, these data can help users better understand the product's experience and make choices.

Although a large amount of evaluation data can help users obtain more detailed information, it has become more difficult for users to find useful information from such a large amount of data. As ordinary users, they often feel disorganized when facing the massive amount of information on the Internet. As a result, there have been some websites and platforms designed to provide users with information to post evaluations. However, while providing evaluation information, this type of website also displays product advertisements and purchase links.

This paper proposes an evaluation retrieval method that uses the commodity value model and implements an evaluation retrieval system based on this method. This evaluation retrieval method uses the special feature of 'different product categories have different evaluation points', and these evaluation points are extracted and combined with the evaluation performance to create a product value model. The evaluation retrieval system uses Twitter data and takes three types of cosmetic products as the research objects, namely foundation, lipstick and cheek blush.

The proposal method is mainly divided into two parts. The first part is the construction of the commodity value model, and the second part is the processing of sorting the retrieved evaluation information according to the needs of users. The specific operation of the first part is divided into the following steps. First, analyze the product characteristics of different product types from the evaluation information on the TF-IDF method network, and perform K-means clustering on these characteristics to obtain the evaluation theme. Then the collected 3000 evaluation information was manually tagged to make a data set for machine learning. Finally, use machine learning to train the data set to obtain the commodity value model.

The second part is achieved by automatically scoring the search results. The scoring standard is the evaluation criterion that users input when searching. The higher the score, the more consistent the piece of information is with the user's search expectations. And the search results are displayed on the front page in the order of the scores from high to low.

Finally, the retrieval system is constructed in the order of data storage layer, data access layer, business logic layer and presentation layer. The data storage layer contains database

and log files. The data access layer contains database management and log management modules. The business logic layer contains commodity value models and information management modules. The Presentation layer includes front-end display pages and back-end systems.

Two evaluation experiments were carried out in this paper. The two experiments evaluated the proposal technique from two aspects. The purpose of the first evaluation experiment is to verify the search efficiency of the proposal method. @cosme and Twitter Keyword searches were selected as two comparison methods. The experiment requires five experimenters to set a search target in advance and then use three search methods in any order to search. During the experiment, the time and the number of clicks for the experimenter to reach the retrieval goal were counted and analyzed. The purpose of the second evaluation experiment is to evaluate the automatic scoring process for search results. The experiment uses MAP (mean average precision) index to evaluate. In order to obtain the MAP value, the experiment randomly selected six search targets set by the experimenter in the first evaluation experiment and used the proposal method to search. The first ten content of the search result is manually judged whether the content is related to the search target, and expressed by 1/0, then calculate the AP value and take the average.

The results of the first evaluation experiment show that the proposal method improves user retrieval efficiency in terms of time and operation. Through the classification of the search target, it can be analyzed that the proposal method is especially suitable for the case of inputting product characteristics to search for the corresponding product. The results of the second evaluation experiment show that the automatic scoring process can sort the data according to user expectations. However, it also showed two problems with the proposal method. The first is the lack of distinction between advertising information and evaluation information, and the second is the lack of data for machine learning.

In future research, the proposed method should solve the problem of the distinction between advertising information and evaluation information and the lack of data for machine learning. In addition, because Twitter has the function of publishing pictures, the search results obtained by analyzing text information are limited. It is expected that the information volume of the retrieval system can be increased by adding the analysis function of the image to the retrieval system.

目次

第1章 はじめに	1
1.1 研究背景	1
1.2 研究目的	2
1.3 本論文の構成	2
第2章 関連研究	3
2.1 緒言	3
2.2 関連知識	3
2.2.1 ウェブクローラー	3
2.2.2 単語ベクトル	5
2.3 関連情報	6
2.3.1 商品評価検索サイトの使用現状	6
2.3.2 Twitter の使用現状	7
2.4 先行研究	9
2.4.1 化粧品レビューの先行研究	9
2.4.2 評価検索の先行研究	10
2.5 本研究の特徴	11
2.6 結言	12

第3章 商品価値モデルを用いた検索手法	13
3.1 緒言	13
3.2 商品価値モデル	13
3.2.1 概要	13
3.2.2 評価トピックと商品特性の抽出	13
3.2.3 学習用データ集の作成	18
3.2.4 ロジスティック回帰分析	21
3.2.5 商品価値モデルの評価	21
3.3 検索システムの構築	24
3.4 結言	28
第4章 評価実験	29
4.1 緒言	29
4.2 実験設定	29
4.3 評価方法	30
4.4 結言	31
第5章 実験結果と考察	32
5.1 緒言	32
5.2 実験結果	32

5.2.1 検索効率を検証する実験	32
5.2.2 スコアリング処理の正確性を検証する実験.....	35
5.3 考察	36
5.4 結言	40
第 6 章 おわりに	41
6.1 まとめ	41
6.2 今後の展望.....	43
参考文献.....	44

図目次

図 1	ウェブクローラーによるデータ収集プロセスのフローチャート.....	5
図 2	日本でのソーシャルメディア使用状況.....	9
図 3	レビュー検索システムの階層図.....	26
図 4	インターフェース.....	27
図 5	タスク達成時間比較図.....	34
図 6	タスク達成クリック数比較図.....	34
図 7	Twitter キーワード検索例.....	37

表目次

表 1 抽出された評価トピックと商品特性（ファンデーション）	16
表 2 抽出された評価トピックと商品特性（リップスティック）	17
表 3 抽出された評価トピックと商品特性（チーク）	18
表 4 タグ付けデータ例	20
表 5 商品価値モデルの評価	23
表 6 評価実験 1 検索タスク内容	32
表 7 実験結果分散分析（時間比較）	33
表 8 実験結果分散分析（クリック数比較）	33
表 9 評価実験 2 検索タスク内容	35
表 10 評価実験 2 結果データ	35
表 11 種類（一）達成時間比較	39
表 12 種類（二）達成時間比較	39

第1章 はじめに

1.1 研究背景

インターネットの急速な発展に伴い、人々のオンライン活動はますます活発になっている。特にオンラインショッピングとソーシャルメディアの利用者が大幅に増加している。近年、オンラインショッピングの世帯利用率は飛躍的に上昇し、個人利用率は全年代平均で7割を超えている（総務省「社会課題解決のための新たな ICT サービス・技術への人々の意識に関する調査研究」平成 27 年）。オンラインショッピングの利点は時間と場所を問わず自由に利用できる便利さに加え、ユーザーの商品評価を参考にできることである。一般的に、オンラインショップにより掲載される商品情報はメーカーの視点に基づいて提供する基本的な商品情報である。そちらの情報は数量が少ない上、ユーザーにとって内容が十分ではない。一方、商品評価はユーザーの視点から、商品の各側面に対する感想などのフィードバック情報である。商品評価機能は実物にアクセスできないというオンラインショッピングの欠点を補うことができる。それは店頭ショップでも使用できない機能である。

商品評価はユーザーにより豊富な商品情報を提供できるし、オンラインショッピングの購買意思決定に役立つ。現在のオンラインショップには多くの種類と数量の商品があり、商品に対応するコメント数も多くある。特に、人気商品のコメント数は数万に達する場合がある。コメントが多いほど、詳細な商品情報をユーザーに提供できるが、ユーザーが商品を全体的に認識するためには時間や手数がかかる^{[1][2]}。多くのオンラインショップはコメント分類機能を提供している、それはユーザーの商品評価によりポジティブ、真ん中、ネガティブの三つの種類に分類している。しかしながら、ユーザーは関心のある商品機能や商品特性に関する情報を効率的に調べることはできない^[3]。

本研究では、商品評価が多いことにより引き起こされる様々な問題を解決す

るために、商品価値モデルを用いる。そのモデルを使った商品評価システムを設計及び実装し、商品評価を感情分類情報や商品特性および機能などにより、直感的と効率的な検索手法を実現する。

1.2 研究目的

本研究ではソーシャルメディアを用いた商品価値モデルにより、ユーザー価値を反映したレビュー検索手法を提案することを研究目的とする。そのため、次の二つの課題を順番に達成することで、本研究の目的を達成する。

課題(1)商品の特徴要素を考慮した商品価値モデルを構築することである。

課題(2)商品価値モデルを用いたレビュー検索システムを開発し、その効果を評価することである。

なお、本研究では、ユーザー嗜好などの感性価値観を考慮するため、メイクアップ商品を対象とし、Twitter データを用いた商品価値モデルによるレビュー検索手法を研究開発する。

1.3 本論文の構成

本論文の構成は以下のとおりである。第一章では本研究の研究背景と研究目的を述べた。第二章の関連研究として関連知識、関連情報、先行研究を述べた後、本研究の特徴を示す。第三章では商品価値モデルを用いた商品評価検索システムの提案手法を説明する。第四章では評価実験の方法を説明する。第五章では評価実験の結果を紹介し、考察する。第六章は本研究のまとめと今後の展望である。

第2章 関連研究

2.1 緒言

本章では、本研究の関連知識、関連研究、先行研究について述べる。最後に本研究の特徴点、つまり本研究と従来研究との違いについて論じる。

2.2 関連知識

2.2.1 ウェブクロラー

本研究では大量のTwitterデータを必要とするツイートに基づいて商品評価を検索する必要がある。手動で収集することは極めて困難である。また、Twitter社が無料で社会に公開しているStandard Search APIは過去7日分までしか検索できないという制約がある。ゆえに、ウェブクロラー技術を利用し、Twitterデータを自動的に取得する。

以下に、クロラーの動作原理とプロセスを紹介する。ウェブクロラーは、プログラムを使用しターゲットウェブサイトへの要求を開始し、ユーザーによるウェブページ情報の閲覧をシミュレートして、ウェブサイトから返された情報を取得する。一般的にネットワーク上のリソースはアクセスリンクに対応し、ブラウザを介してアクセスできるため、ブラウザはデータ要求の受信と分析および表示を担当する。ウェブクロラーは情報をクロールするだけでよく、ユーザーにページ表示を提供する必要はなく、ブラウザの代わりにターゲットリンクへの要求を直接行える。ウェブクロラーは、返されたデータを解析しフィルタリングするか、ユーザーが関心のあるデータを直接取得するか、などの処理を実行してデータ保存できる。その結果、検索システムはこのデータを利用できるようになる。

本研究のウェブクローラーは、Twitterの検索コンテンツとユーザーのホームページをクロールする機能である。Seleniumにより実現され、3000件以下のツイートをクローラーのプロジェクトに応用する。Twitterの高度な検索と組み合わせると、特定の言語、特定の期間などでTwitterコンテンツをフィルタリングするという目的を達成できる。本研究に使うウェブクローラーはセレンとTwitterの高度な検索を組み合わせ構成とデータのクロールを容易にすることを目的としており、小規模または予備テストに適応する。ウェブクローラーによりTwitterデータを収集するプロセスは次の通りである：

1. クロールデータの関連パラメータを設定する：検索キーワード、クロールしているページ数、クロール言語は‘ja’、ブロック画像を設定、URL を設定する。
2. ブラウザーを模擬する。
3. ツイートを発表時間順に獲得する。
4. ページ数は設定ページ数より大きい場合はクロールを終了する、それ以外の場合は手順3に戻る。
5. ツイート、発表時間、コメント数、Like 数、リツイート数をデータベースに保存する。

図1にウェブクローラーのプロセスを説明するフローチャートを示す。

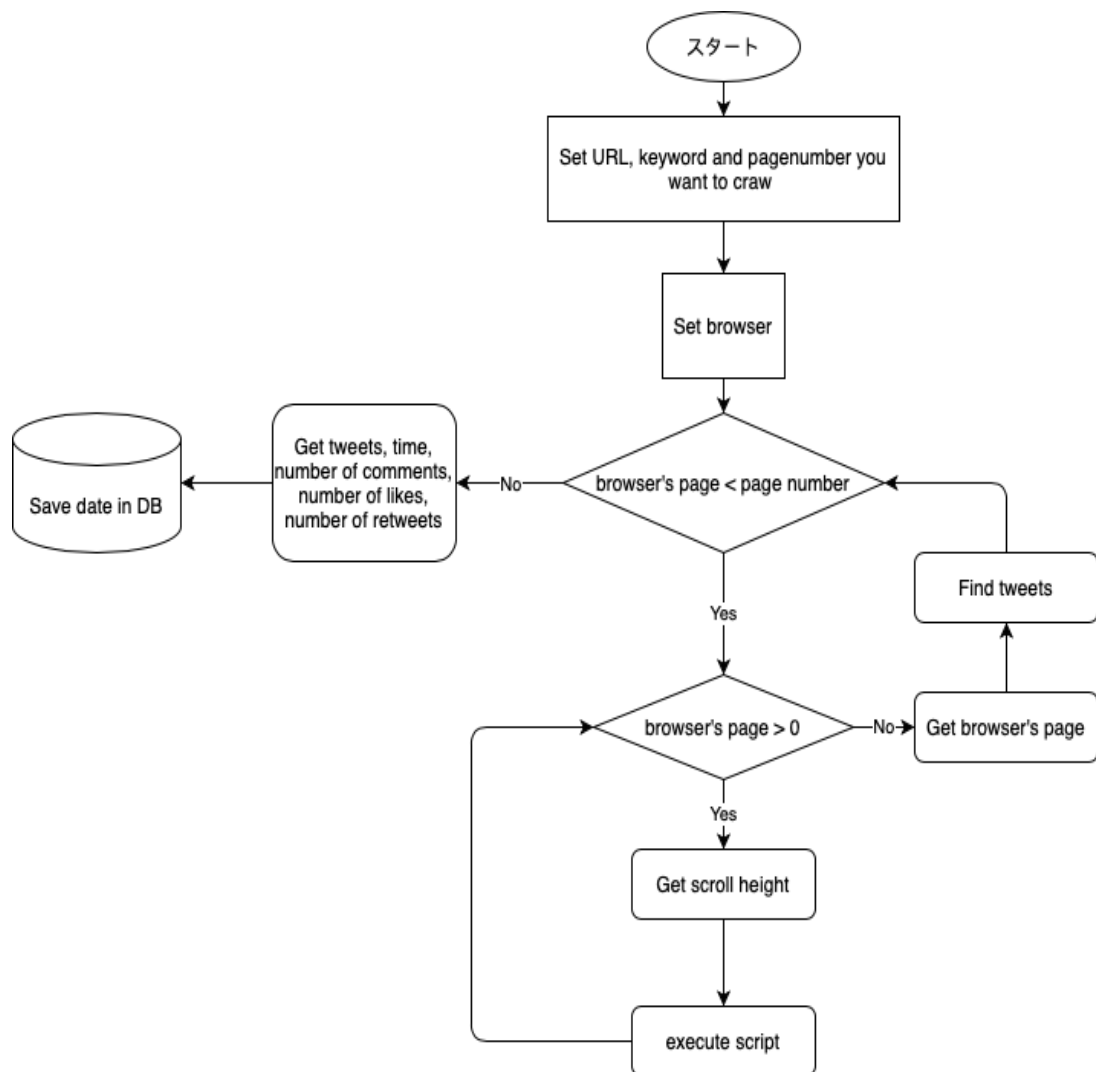


図 1 ウェブクローラーによるデータ収集プロセスのフローチャート

2.2.2 単語ベクトル

自然言語処理に機械学習アルゴリズムを使用するには、テキストからデジタル化された特徴を抽出する必要がある。単語ベクトルを使用し単語を表し、テキストの特徴をデジタル化できる。より一般的な単語ベクトル表現には二つのタイプがある。

一つは従来のone-hotベクトルであり、もう一つは分散型単語ベクトルである。

最初、コーパスに含まれるすべての単語に番号が付けられる。単語の単語ベク

トルは、対応する番号位置の実数が1、他の位置の実数が0である。単語ベクトルの次元は、コーパスの対応する辞書サイズである。このタイプの単語ベクトル生成方法は単純で理解しやすいが、次の問題点が存在する。

1. 変数の数が多くなり過ぎると、計算に時間がかかったり、保存問題が発生したりする可能性が大きい。このような場合には、低次元のベクトルで表現することが考えられる。

2. 単語のセマンティクスとコンテキスト情報の欠如が問題となる。同義語間の同様の意味関係は、0-1ベクトルの形で表現することはできない。同時に、0-1表現は、コーパス内の単語と他のいくつかの単語で構成されるコンテキスト情報を失う。

2.3 関連情報

2.3.1 商品評価検索サイトの使用現状

現在、アットコスメ^[4]をはじめ広く使われている化粧品レビュー検索サイトが存在する。2020年4月時点においてアットコスメのユーザー数は3,670,000人を超えている(デバイス:PC+スマートフォン)。アットコスメ公式サイトにおけるランキング、ブランド、ブログ、お買い物、口コミなどの機能を使用できる。またアットコスメではオンラインとオフラインを組み合わせた次のようなサービスを提供している。

(1) アットコスメ・アットコスメショッピング・アットコスメストア間のシームレスなユーザー体験

アットコスメにおける口コミ・コンテンツ閲覧から、「アットコスメショッピング」「アットコスメストア」における購買を行え、ユーザーが一貫して利用しなくなるような体験づくりを提供している。

(2) アットコスメ価値を用いたコンテンツ

アットコスメの編集力を生かしたタイアップをはじめとしたコンテンツを発信している。また、JR 原宿駅前に 2019 年にオープンしたアットコスメストアのフラッグショップ「アットコスメ TOKYO」を活用したユーザー向けイベントなどを充実させている。

アットコスメは日本最大のコスメ・美容に総合サイトであるが、商品レビューを検索したいユーザーにとって、アットコスメを利用するのは最優選択とは言えない。アットコスメには 16,344,000 件以上の口コミが掲載されている(2020 年 1 月まで)。研究背景で述べたとおり、口コミが多いほど、詳細な商品情報をユーザーに提供できるが、ユーザーとして有用な情報を入手するまで手数が増えている。なお、アットコスメはブランドとの連結が行なっており、さらにアットコスメショッピングとアットコスメストアでコスメ商品を販売する。だから、実際より良い評価をもらうためにユーザーにポイントあるいはプレゼントの贈与行為を通じて、コメントに偏りが出してしまう可能性がある。また、商品販売を促進するために、良い評価と悪い評価を戦略的に展示することも可能である。良い評価を見やすい場所に置いたり、注目された字体と色を使ったりすることが望まれる。

2.3.2 Twitter の使用現状

ソーシャルメディアとは、www のような情報発信技術を用いて、社会的関係を利用するために設計されたメディアである。双方向のコミュニケーションができることは特長である。現在、影響力を持ち始めたメディアは、LINE、Facebook、Twitter や Instagram などの個人間の情報発信が可視化されやすくなるメディアである。ソーシャルメディアはインターネットを前提として文字情報、画像、映像などのコンテンツを手軽に大衆に伝えることができる。その結果、人々をコンテンツの消費者から生産者へと変える。

日本での各ソーシャルメディアの使用状況を図 1 に示す。自ら情報発信や発言を積極的に行なっている人の比率に注目すると、上から一番目の LINE は 17.0%であり、次の Twitter は 7.7%であり、Facebook は 5.3%であり、ブログは 4.6%である（総務省「ICT によるインクルージョンの実現に関する調査研究」2018¹）。LINE での積極的発信や発言する人は Twitter より多いが、ソーシャルメディアの種類について考えれば、LINE はメッセージチャット型であり、Twitter は交流系拡散型 SNS である。交流系拡散型 SNS はメッセージチャット型より情報を大衆に広める効果が予想できる。

Twitter は 140 字以内の短い文章が投稿できるソーシャルメディアであり、画像や短い動画の投稿もできる。140 字以内という制限により、情報の発信と受信は手軽に行うことが可能になる。また、Twitter は投稿がリアルタイムで流れていく特徴がある。かつ、Twitter は拡散性の高いメディアであるため、フォロワーの枠を超えて情報が拡散していくことも大きな特徴である。一般的に、ユーザーが Twitter に投稿するツイートは自発的な発話行為であるに対し、公式サイトや比較サイトなどで掲載された商品レビューは「誰かにインタビューされた」という雰囲気がある。そのシーンの下に、ユーザーは積極的に商品进行评估する意欲がないが、他人に聞かれると良い評価の返事をする事になりやすい。それに対して、Twitter に投稿する商品関連情報を含むツイートはより本音に近い発言と考えられる。

¹ <https://www.soumu.go.jp/johotsusintokei/whitepaper>

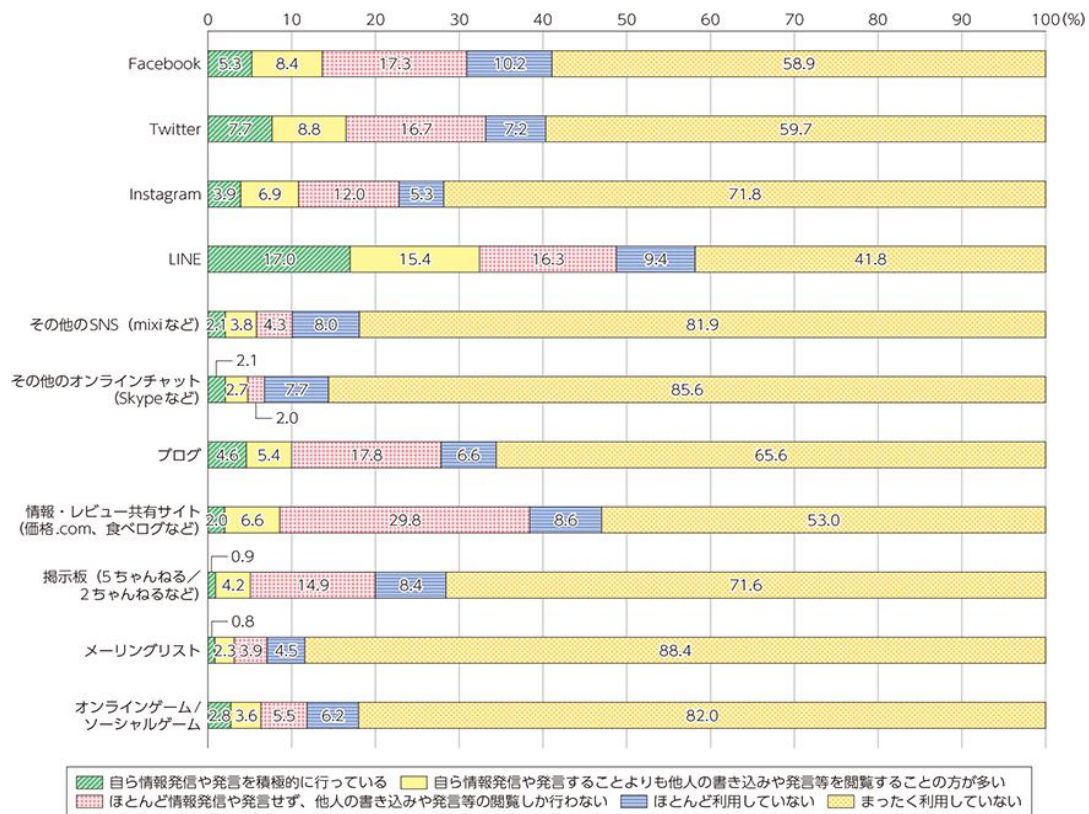


図 2 日本でのソーシャルメディア使用状況

(<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/image/n4202010.png>)

したがって、本研究では提案した検索手法は Twitter データを利用する。Twitter データの 140 字以内制限による「随意性」とリアルタイムによる「即時性」という二つの特徴を活用し、ユーザーにとって信頼度が高い情報を提供するレビュー検索システムを目指す。

2.4 先行研究

2.4.1 化粧品レビューの先行研究

松波ら^[5]は、化粧水に類似した化粧品カテゴリーや化粧品とは異なる化粧品カテゴリーのレビューにおける評価表現の傾向を確認し、各化粧品カテゴリーに対応する表現辞書の構築方法を検討している。

酒井ら^[6]は、美容レビューで一般的な単語を抽出し、レビュー文の各カテゴリに現れる一般的な単語の数を使用し、余弦の類似性を計算および比較している。さらに、余弦類似度が計算され、頻繁な単語の有無はそれぞれ 0 と 1 としている。これらの比較の結果に基づく、化粧品カテゴリにより共有可能な評価項目があることを確認している。

白田ら^[7]は、アットコスメのマスカラのレビュー文を分析している。レビュー文から名詞、形容詞、動詞を抽出し、RIDF（残差 idf, residual inverse document frequency）を使用して重みを計算し、単語の共起と頻度の情報を使用し、コンテンツに基づいたグラフを生成している。さらに、レビューごとにグラフを生成している。グラフ可視化ソフトウェアである gephi を拡張した Consumer Behavior Analyzer を用いて話題ごとに抽出している。そして、抽出した話題から、マスカラのレビューではマスカラの持続力や落とす際の特徴が記述されている。

以上の先行研究では、化粧品カテゴリにおける評価項目は共有項目と特有項目があることを示唆した。しかし、化粧品カテゴリ毎に特有評価項目を活用する消費者にレビュー情報を提供する研究が欠けている。

2.4.2 評価検索の先行研究

立石ら^[8]は、ウェブ文書の意見を取り扱うことをタスクとし、次のような機能をもつ評判情報検索システムを開発した。①特定商品に関する意見を検出できる、②意見に該当する箇所を抽出できる、③抽出した意見を肯定・否定に分類できる。

二本木ら^[9]は、＜対象＞＜評価＞の構造を利用し、評価に関する表現を抽出し定量化し、効率的な分析・検索手法を提案している。レビュー検索サイトに投稿されていた大量の自由記述文から、＜評価＞の辞書と品詞情報を利用し、＜対象＞と＜評価＞の抽出は商品の種類に関わらず 0.7 の再現率が見られた。

姚ら^[10]は、消費者がインターネット上に投稿された口コミを信用する場合、

信頼度の違いと購買行動への影響を検証するために 3 つの仮説をたてている、一つ目は消費者が購入したいと考えていた商品の良い口コミはその商品の購買を促進し、購入を検討していなかった商品への悪い口コミはその商品の購買を阻害する。二つ目は詳細データを書き込んだ口コミは、客観的な口コミより信用できる。三つ目はアドバイスを書き込んだ口コミは、使い心地や個人的な感想・印象を書き込んだレビューより、信用できる。彼らはインターネットに掲載する口コミは消費者の購買行為に影響を与えることを検証した。

Pham ら^[11]は、least square method を用いてホテルのレビューを対象とし、評価項目毎にスコアを推測する手法と最も重要な評価項目を推測する手法を提案している。この手法では、レビュー中に評価項目に関する内容がない場合でも、その評価点を推測している。一方、Zhang ら^[12]は、ユーザーのレビューテキストに対して、評価項目毎にユーザーの評価点を推測する手法を提案している。提案手法では、まず評価項目に関する文をレビューテキストから抽出し、その文から評価点を推測する点、また評価項目に関する文がレビューテキストに存在しない時は評価点の推測を行わない点に特徴がある。

2.5 本研究の特徴

インターネットに掲載する大量の商品評価に注目する研究が多数存在し、レビュー検索手法もいくつか提案されている。従来の研究では商品評価を肯定表現と否定表現の出現頻度によりポジティブとネガティブな評価として捉える研究は多いが、商品カテゴリー毎に持っている特徴評価項目に関する表現とレビュー検索との関係については検討されてきていない。

本研究では、商品特徴評価項目とユーザーの価値判断基準を考慮した商品価値モデルによるレビュー検索手法を提案している。ファンデーション、リップスティックとチークの三つのメイクアップ商品のカテゴリーを設定する。カテゴリー別で商品特性を抽出し、商品特性をクラスタリングすることで評価トピ

ックを分類する。抽出した商品特性に基づいて商品価値モデルを構築する。従来のレビュー検索手法は主にキーワードと一致するワードが含まれたデータを検索結果とする。一方、本研究の特徴はユーザーの価値判断基準に基づいて、ユーザーが関心のある情報を含むレビュー文を優先して表示する点にある。

2.6 結言

本章では、提案手法に関連する基礎知識及び関連情報について述べた。次には従来の研究をまとめて紹介した。最後には本研究と従来研究との違い点を説明した。

第3章 商品価値モデルを用いた検索手法

3.1 緒言

本章では、Twitter データを用いた商品価値モデルによるレビュー検索手法を紹介する。

3.2 商品価値モデル

3.2.1 概要

本研究では、商品価値モデルを用いたレビュー検索手法を提案する。最初、商品価値モデルというものを紹介する。ユーザーがあるアイテムのレビューを検索する際、そのユーザーは多少の独自の判断基準を持っていると言える。ここでは、ユーザーの価値判断基準と呼ぶ。ユーザーの価値判断基準を考慮したモデルを、本研究では商品価値モデルと呼ぶ。本研究は複数のメイクアップ商品カテゴリーの中、日常的広く使われているファンデーション、リップスティック、チークを対象としている。異なる商品アイテムは自らの商品特性を持っている。例えば、リップスティック商品进行评估する際、商品の色を中心に表現するものが多いことに対し、ファンデーション商品ではカバー力が多く評価される。本研究ではインターネットに既存のレビューの中よく評価された項目をユーザー視点から考えた重要な商品特性と見なす。ユーザーが本レビュー検索システムを使用する際、ユーザーから入力された検索タスクを商品価値モデルにより分析してから対応する商品レビューを効率的に提供できるようにする。

3.2.2 評価トピックと商品特性の抽出

本研究では、カテゴリー別でデータ TF-IDF 値ランキングにより、上位 20 個の形態素を商品特性候補とする。そして、著者は商品特性候補を意味判断して重複なものを削減する。例えば、商品特性候補に「潤い」と「うるおい」がある場合、一つだけ保存する。

収集したデータはファンデーション、リップスティック、チークという三つのカテゴリーにおける各 18,000 件である。データ期間は 2020 年 6 月 1 日から 2020 年 11 月 30 日までである。データ一件を一文書と見なし、形態素解析を行い、品詞ごとに解析した。名詞、固有名詞、形容詞、形容動詞という四つの品詞を抽出する。形態素解析のため用いた日本語辞書は MeCab である。それから、TF-IDF 法により商品特徴を特定する。TF-IDF 法は単語の重要度を示す指標である。TF-IDF (t, d) は高い場合は単語 t が文書 d の特徴後語と言える。具体的には、次の式 (3) ～ (5) により TF-IDF 値を計算する。

$$TF-IDF_{i,j} = tf_{i,j} * idf_i \quad (3)$$

$$tf_{i,j} = \frac{n_{i,j}}{\sum_s n_{s,j}} \quad (4)$$

$$idf_i = \log \frac{|D|}{|\{d: d \in t_i\}|} \quad (5)$$

$n_{i,j}$ 文章 j の中単語 i の出現頻度

$\sum_s n_{s,j}$ 文章 j 内のすべての単語出現頻度

$|D|$ 文書総数 $|\{d: d \in t_i\}|$ 単語 i を含む文書数

TF 値は単語の出現頻度であり、IDF 値は単語の逆文書頻度である。

以上の、公式で求めた TF-IDF 値により形態素をランキングし、上位 20 語を商品特性候補とする。著者は商品特性候補の内容を確認し、商品特性を判断したデータを商品特性として抽出する。

最後は処理した商品特性をクラスタリングし、評価トピックを抽出する。クラスタリング手法は K-means 法を使用する。K-means 法は、非階層的クラスタリングと呼ばれる分類手法の一種である。教師なし学習であるため、事前にトレーニングデータを提供する必要がなく、評価関数を使用してターゲットを分類できる。K-means 法では、各クラスターの重心は、クラスターに属する全部の要素の平均を計算することにより計算され、次にこれらの要素は、最も近い重心を持つ

クラスターに再分類される。繰り返しても変化が発生しない場合は、収束しているとみなす。

次の表 1、2、3 で示しているのは TF-IDF 法と K-means 法により抽出した商品トピックと商品特性である。

商品アイテム	評価トピック	商品特性
ファンデーション	ブランド	日本系
		韓国系
		中国系
		欧米系
		他
	ファンデーション種類	パウダーファンデ
		クッションファンデ
		クリームファンデ
		リキッドファンデ
		他
	評価表現	効果良い
		効果悪い
		カバー力強い
		カバー力弱い
		使用感良い
		使用感悪い
		マット
		興味ある
		安い
		他
	適用季節	春
		夏
		秋
		冬
		他

表 1 抽出された評価トピックと商品特性（ファンデーション）

商品アイテム	評価トピック	商品特性
リップスティック	ブランド	日本系
		欧米系
		韓国系
		中国系
		他
	色系	ピンク系
		ヌード系
		オレンジ系
		ブラウン系
		レッド系
		他
	評価表現	落ちにくい
		落ちやすい
		効果良い
		効果悪い
		使用感良い
		使用感悪い
		マット
		人気
		興味ある
		他
	適用季節	春
		夏
		秋
		冬
		他

表 2 抽出された評価トピックと商品特性（リップスティック）

商品アイテム	評価トピック	商品特性
チーク	ブランド	日本系
		欧米系
		韓国系
		中国系
		他
	色系	ブラウン系
		ピンク系
		オレンジ系
		パープル系
		レッド系
		コーラル系
		他
	評価表現	効果良い
		効果悪い
		使用感良い
		使用感悪い
		色味良い
		色味悪い
		マット
		人気
		興味ある
		他
	適用季節	春
		夏
		秋
		冬
		他

表 3 抽出された評価トピックと商品特性（チーク）

3.2.3 学習用データ集の作成

商品価値モデルは教師あり学習を用いて作成するため、学習用データが必要である。教師あり学習では、正解データ（目的変数）を含むデータセットを入力として利用する。目的変数を除く残りのデータ（説明変数）から得られる出力結果に着目する。その値ができるだけ正解に近くなる様な特徴量を探し出し、モデルを作成する。最後に新規データに作成したモデルを適用し、予測結果を得る。

学習によって得られる推測モデルの性能には学習用データの量と質が影響する。データの量の重要性については実験により確かめられているが、データの質については、誤りが過多のデータで学習しても、十分な精度を出すことは困難である。本研究では、学習用データを人手で作成することを著者一人で行っている。研究対象アイテムのメイクアップ商品（ファンデーション、リップスティック、チーク）をキーワードとする検索された 3,000 件ツイートに人手でタグを付ける。タグ内容は表 1 から表 3 で示した評価トピックごとの商品特性である。

タグ付けのデータ例を表 4 に示す。

ツイート		ブランド タグ	季節 タグ	評価 表現	色系
1	パリにある Rouge のコスメ、使い勝手も良さそうだし、パケも高級感あって素敵だな……リップパレットの色合いがとても好き	欧米系	秋、冬	効果良い	ブラウン系
2	インテグレート グレイシィ リップスティック 31 番♡ ウルウル、ツヤツヤ、リップクリームみたいに軽くて唇本来の色のように淡く色づくんだって。恋が叶う「恋コスメ」って言われているんだって♡	日本系	春、夏	効果良い、使用感良い	ピンク系
3	ネクストブレイク間違いなし！ 韓国コスメの『ピーター』のリップスティック 価格以上の価値が感じられるコスメとして大人気 ふんわりとしたセミマットで、重ね塗りして鮮やかな色味も、ポンポンつけて淡い発色もどちらも◎ 見た目も可愛い #韓国コスメ #リップ https://news.merumo.ne.jp/article/genre/10178069 ...	韓国系	他	マツト、人気	ピンク系
4	【キャンメイク】ステイオンバームルージュ。リップバームの保湿力・口紅の発色・グロスのツヤ・UVケアすべてを兼ね備えた発色リップスティック→発色がとてもいいし、リップクリームを塗らなくても唇があれったり乾燥したり今のところありません	日本系	秋、冬	使用感良い、効果良い	他

表 4 タグ付けデータ例

3.2.4 ロジスティック回帰分析

商品価値モデルを構築する際、ロジスティック回帰分析により多種類を分別する、具体的には over-vs-all アルゴリズムを使用した。商品特性の季節を例として説明する。季節という商品特性において四つの類があり、ここでは $y=1, y=2, y=3, y=4$ で表す。学習用データ集を使用し、四つの二元分類問題に分けた。類 1 を区別するため、類 2、3、4 をネガティブクラスとし、類 1 をポジティブクラスとする仮データ集を作り、適切な分類器を取り付ける。このモデルを $h_{\theta}^{(1)}(x)$ とする。次に、類 2 をポジティブクラスとし、類 1、3、4 をネガティブクラスとする。このモデルを $h_{\theta}^{(2)}(x)$ それから、類 3 と類 4 をポジティブクラスとし、モデルを $h_{\theta}^{(3)}(x)$ と $h_{\theta}^{(4)}(x)$ とする。そして、 $h_{\theta}^{(i)}(x) = p(y = i|x; \theta)$ $i = (1, 2, 3, 4)$ が得られる。最後に、予測する際に、全ての分類器をランニングし、入力変数に対して一番可能性が高い出力変数を選ぶ。ロジスティック分類器 $h_{\theta}^{(i)}(x)$ をトレーニングし、 i は全ての可能性を持つ $y=i$ 。それから、新しい x 値を入力し、それを用いて予測する際、全ての分類器に x を入力し、 $\max h_{\theta}^{(i)}(x)$ が得られる。

3.2.5 商品価値モデルの評価

商品価値モデルを評価するため、三つの研究対象カテゴリーは各 150 件、合計 450 件に対し、四つの評価トピックごとに、商品レビューとみなせるテキスト

トを人手でタグづけた。これを正解とみなしたときの商品価値モデルの精度、再現率、F 値を測った。結果を表 5 に示す。

商品カテゴリー	評価トピック	適合率	再現率	F 値
ファンデーション	ブランド	91.39	73.62	81.55
	季節	97.56	96.34	96.95
	種類	90.28	77.64	83.48
	評価表現	85.13	86.91	86.01
リップスティック	ブランド	94.34	83.62	88.66
	季節	96.73	97.24	96.98
	色系	92.01	73.59	81.78
	評価表現	83.4	76.27	79.68
チーク	ブランド	92.08	75.84	83.17
	季節	93.24	86.58	89.79
	色系	88.94	81.37	84.99
	評価表現	80.62	63.68	71.16

表 5 商品価値モデルの評価

適合率について、「評価表現」は全て 90%未満であり、チークの「色系」は 88.94%であり、それ以外は全部 90%を超えた。一方、三つのカテゴリーの「季

節」再現率は 63.68%から 86.91%であった。「季節」は 86.58%から 97.24%であった。「評価表現」についての評価結果は改善の余地が示した。特に、チークの「評価表現」の再現率は低い。その理由は、チークに関するレビューでは、評価表現として言及された単語が「発色」以外のものは滅多に見られない。

3.3 検索システムの構築

Twitter データを用いた商品価値モデルによりレビュー検索システムの実現には次の手順で行なう。

STEP1：ウェブクローラーを使用して Twitter データを収集する

STEP2：既存の化粧品レビューから研究対象（ファンデーション、リップスティック、チーク）の評価トピックおよび商品特性を抽出する

STEP3：人手でタグを付けて学習用データ集（3,000 件）を作成する

STEP4：商品価値モデルを教師あり機械学習により構築する

STEP5：構築した商品価値モデルを用いて全てのデータにタグ付けを行う

STEP6：検索結果の自動スコアリング処理を実現する

STEP7：インターフェースを設計し、レビュー検索システムを作成する

レビュー検索システムは図 3 に示す三層 Client Server System を発展させた四層のシステムとして構築した。下からデータストレージ層、データアクセス層、ビジネスロジック層、およびプレゼンテーション層である。各層の詳細な導入は次に示している。

1、データストレージ層は、システムに関連するデータの永続的ストレージを実現するために使用される。このストレージ層には、データベースとログファイルが含まれる。

2、データアクセス層は、データストレージ層へのデータアクセス機能を実装するために使用される。データベース管理とログ管理を含む。データベース管理では、ストレージ層でのデータベース・アクセス機能を実現する。ログ管理では、システムのキー操作とエラー情報をログファイルに記録する。

3、ビジネスロジック層は、レビュー検索システムのビジネスロジックを実装するために使用され、データアクセス層機能を使用している。この階層には、商品価値モデル・モジュールと情報管理が含まれる。

4、プレゼンテーション層は、レビュー検索システムのインタフェースである。この層には、レビュー検索システムのフロントエンド・システム情報表示インタフェースと、バックエンド・システムの関連する管理インタフェースが含まれる。

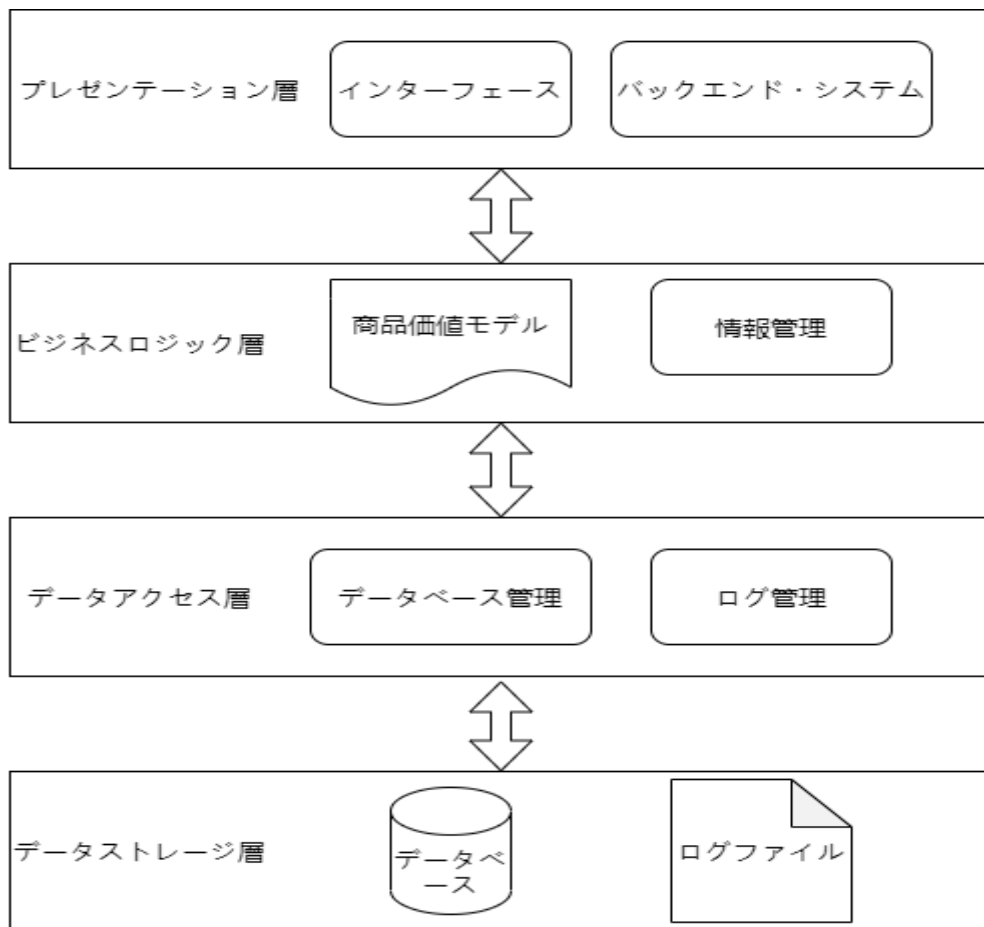


図 3 レビュー検索システムの階層図

検索システムのインターフェースを図4に示す。ユーザーは左側の選択肢から関心のある商品特性を選択し、検索ボタンを押す。それから、右側に検索結果をスコアリングした結果を順番に表示する。

ファンデーション

リップスティック/口紅

チーク

1.ブランド

☒ 日本系
☐ 韓国系
☐ 中国系
☐ 欧米系
☐ 他

2.種類

☒ パウダーファンデ
☐ クッションファンデ
☐ クリームファンデ
☐ リキッドファンデ
☐ 他

3.季節

☒ 春
☐ 夏
☐ 秋
☐ 冬
☐ 他

検索

図 4 インターフェース

comment	date	like	name	retweet
私アルビオンのスキンケア用品は使用しても化粧品はサンブル使う程度なものでww でもアルビオンのファンデもおしろいもめちゃくちゃ良いのは使った事あるので分かります!! 肌に自然に馴染むし塗ってる感なくて本当に素晴らしいと思います!とお伝え下さいましあ(´_ゝ´)!!!	2020-11-08T13:36:47	0.0	爆柱の Ra Ra Ra * R E!...*	1.0
今のところ自分に合うベース【春夏】①プリマヴィスタ下地②ミシャククッションファンデ(マット)③プリマヴィスタ下地④チャコットパウダー-or kissパウダー⑤ポルジョ下地⑥ポルジョパウダー【秋冬】①ポルジョ下地②ポルジョパウダー③kiss下地④kiss リキッドファンデ	2020-11-08T03:30:15	0.0	りうす	5.0
MUFのファンデ、カバー力=厚まいたいという固定概念を覆すような感触です。以前仰っていたように毛の柔らかいブラシでつけていただいても密着性いいのでピタッと肌に付くのにナチュラルな仕上がりには驚きます。冬にこそ本領発揮するファンデです♡	2020-11-08T07:31:13	0.0	手を洗う。こあら0	0.0
いや、見習うだなんてファンデはオルビスとセザンヌ(ドラッグストアで手に入る端安モノ)使ってます私も30代入るまでほとんどファンデ使ったことなくこの頃親しいママ友に色々教えてもらって習得してるんですがやりだすと楽しくて色々欲しくなっちゃいますよー	2020-11-08T13:35:08	0.0	なな	0.0
ファンデじゃないの? ってレベルで肌をキレイに魅せる下地、ポルジョーのモイスチュアライジングファンデーション...				

3.4 結言

本章では、本研究で提案する商品モデルによりレビュー検索手法を商品価値モデルの作成方法を紹介した。実際の検索システムを実現する手順で説明した。

第4章 評価実験

4.1 緒言

本章では、提案手法の評価実験について述べる。4.1 節では実験の設定について説明する。4.2 節では実験の評価方法を紹介する。

4.2 実験設定

提案手法の検索効率とスコアリング処理の正確性を評価するために、二つの評価実験を設定する。

評価実験(1) 検索効率を検証する実験

実験目的は商品価値モデルを用いたレビュー検索手法はレビューを効率的に検索できるかを検証することである。実験参加者は事前に検索タスクを設定し、提案手法と比較手法を使用したレビュー検索を行う。

提案手法：Twitter データを用いた商品価値モデルによるレビュー検索手法

比較手法①：アットコスメ公式サイトにおける口コミ検索

比較手法②：Twitter 公式サイトにおけるキーワード検索

実験参加者は予め検索タスクを設定する。ファンデーション、リップスティックとチークという三つのカテゴリーに関する検索タスクを各 1 件、提案手法と二つの比較手法でレビュー検索を行う。参考になるレビュー文が現れる次第、実験が終了する。終了するまでの目標タスクの達成時間および達成クリック回数を数えて比較する。検索手法の使用順序に影響を受けることを控えるために、実験参加者は検索手法を使用する順序はランダムとする。

実験前、実験参加者に実験事項を説明する。具体的には以下である。

「予め検索タスクを設定してください。つまり、今回のレビュー検索を通し、どんな情報が知りたいですか。閲覧した情報は十分に参考になると感じた場合、検索タスク達成すると報告してください。十分な情報を参考とするために、参考になる情報は幾つか閲覧しても大丈夫です。検索手法の使用順序はランダムです。」

評価実験(2) スコアリング処理の正確性を検証する実験

実験目的は評価実験が検索したレビュー順序の妥当性を評価することである。

異なる検索目標を提案手法で検索する六つの検索タスクを設定する。検索タスクは評価実験(1)で実験参加者に検索された検索タスクの中からランダムに抽出し、三つの商品カテゴリーごとに二つのタスクを抽出する。各ケースの順序前10枚のレビューを手で内容を確認し、検索目標と関連するレビューに「1」をつけ、関連しないレビューに「0」をつける。手で0/4タグを付与する際、著者と本学の学生二人合計三人で判定し、多数の人が判定した結果を最終結果とする。

4.3 評価方法

平均値はデータセットの中間的な値であり、標準偏差は、すべてのデータの分散度を表す絶対値である。提案手法と比較手法との間有意差ありを検証するためにANOVA（分散分析）とt検定により評価する。検索タスクは異なる種類あるため、全体的な差異を検証してから、さらに種類別にt検定を行う。

機械学習、自然言語処理、情報検索などの分野において、一般的には使用する評価指標は正確率、適合率、再現率およびF1-測定である。しかしながら、IRのグラウンドトゥルスは、一般的に順序付きリストであり、Boolタイプの順序なし

コレクションではない。それらがすべて見つかった場合では、3 位と 4 位の区別はそれほど大きい意味を持ってないため、1 位と 100 位の区別を重視しなければならない。だから評価実験(2)では MAP(Mean Average Precision)を評価指標として用いる。MAP は次の式で求められる。

$$AP = \int_0^1 P_{(r)} dr$$

$$MAP = \frac{1}{m} \sum_{i=1}^m AP_i$$

4.4 結言

本章では、提案手法を評価するため二つの評価実験を紹介している。提案した商品価値モデルを評価するためのレビュー検索実験とスコアリング処理の正確性の検証実験について述べた。

第5章 実験結果と考察

5.1 緒言

本章では提案した検索手法の評価実験結果を示し、考察する。

5.2 実験結果

5.2.1 検索効率を検証する実験

実験参加者は事前に設定した評価タスク内容は表 6 に示す。

商品種類	評価タスク番号	検索タスク内容
ファンデーション	1	冬に使うファンデーション
	2	シャネルとディオルのクッションファンデ
	3	乾燥しやすい肌を潤う
	4	カバー力強いクッションファンデ
	5	SUQQU のクリームファンデの効果
リップスティック	6	キャンメイク口紅の落ち持ち
	7	オレンジ系リップスティックの推薦
	8	YSL の人気リップスティック
	9	落ちにくい口紅
	10	秋色口紅
チーク	11	韓国産安いチーク
	12	推薦人気があるオレンジ系チーク
	13	チークの塗る範囲
	14	血色感チーク
	15	SUQQU、セザンヌチークの発色

表 6 評価実験 1 検索タスク内容

実験中、実験参加者は各検索タスクを達成するまでの時間（秒）とクリック数を計数した。達成する時間について、提案手法は平均 16.93 秒であり、平均 109.20 秒の比較手法 1 と平均 71.33 秒の比較手法 2 より検索時間を短縮した。提案手法を二つの比較手法と各自比較するために、t 検定を行った。p 値は 0.001 と 0.034 であるため（表 6）、検索時間には提案手法と比較手法 1、比較手法 2 の間に有意差を示した。提案手法は比較手法より短期間で検索タスクを達成できることがわかった。

	平均値(s)	標準偏差	平均値の差	t 値	p 値
提案手法	16.93	16.98	-94.27	-3.996	0.001**
比較手法 1	109.20	88.59			
提案手法	16.93	16.98	-56.40	-2.346	0.034*
比較手法 2	71.33	99.39			

p < 0.05* p < 0.01**

表 7 実験結果分散分析（時間比較）

達成するクリック数について、提案手法は平均 3.87 回であり、比較手法 1 は 18.87 回であり、比較手法 2 は 7.47 回である。提案手法を二つの比較手法と各自比較するために、t 検定を行った。p 値は 0.015 と 0.008 であるため（表 8）、比較手法と比べ、提案手法は便利な操作で検索タスクを達成することを示した。

	平均値(s)	標準偏差	平均値の差	t 値	p 値
提案手法	3.87	4.09	-15.00	-2.757	0.015*
比較手法 1	18.87	20.35			
提案手法	3.87	4.09	-3.60	-3.76	0.008**
比較手法 2	7.47	6.48			

表 8 実験結果分散分析（クリック数比較） p < 0.05* p < 0.01**

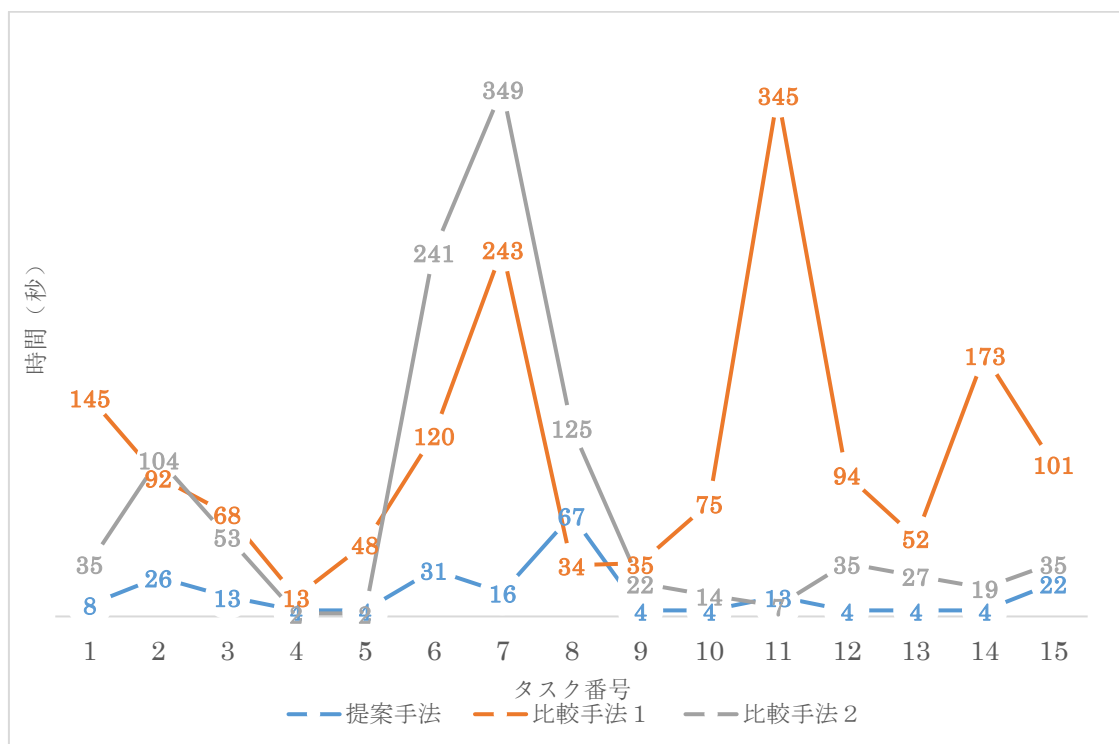


図 5 タスク達成時間比較図

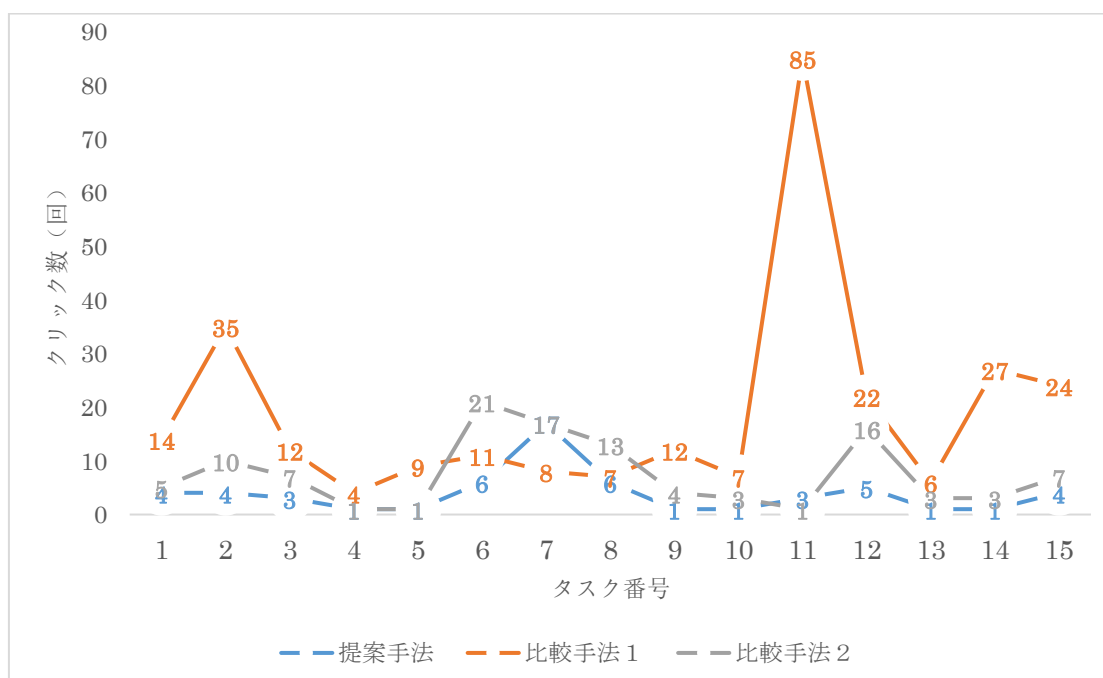


図 6 タスク達成クリック数比較図

実験結果の達成時間と達成クリック数の詳細数値を図5と図6に示す。

図5と図6に示しているように、提案手法は全体的に比較手法より曲線が穏やかであることに對し、比較手法1と比較手法2は極端な数値が幾つか存在する。それは検索タスク7と11であり、内容は「オレンジ系リップスティックの推薦」と「人気があるオレンジ系チーク」である。二つの検索タスクは異なる商品カテゴリーに属するが、二つとも商品特性を述べて対応商品を検索するケースである。ゆえに、検索タスクを分類し考察する必要があると考える。

5.2.2 スコアリング処理の正確性を検証する実験

AP 値は 0.387 から 0.883 までである。MAP は AP の平均値であるので、0.641 である。

	ケース	相関性 (1:あり 0:なし)	AP
ファンデーション	1	0/1/0/1/1/1/0/0/1/0	0.564
	2	0/1/0/0/0/0/1/0/1/0	0.387
リップスティック	3	1/1/0/1/1/1/0/0/1/0	0.821
	4	1/0/1/0/0/0/0/1/0/0	0.680
チーク	5	0/1/0/1/1/0/0/0/1/1	0.509
	6	1/1/1/0/1/1/0/0/1/0	0.883

表 9 評価実験2 検索タスク内容

評価実験1の検索タスク内容(5.2.1 表6を参照)からケースを六つ抽出した。ケース1～6の内容を表9に示す。

ケース番号	検索タスク内容
1	冬に使うファンデーション
2	シャネルとディオールのクッションファンデ
3	キャンメイク口紅の落ち持ち
4	オレンジ系リップスティックの推薦
5	韓国産安いチーク
6	人気があるオレンジ系チーク

表 10 評価実験2 結果データ

本実験の相関性判定は人手で行ったが、その際の宣伝広告はすべて相関性なしと判定している。そのため、広告ツイートを先頭に表す場合は低い AP 値になる、その例はケース 4 である。AP 値が高いケースは 3 と 6 がある。ケース 3 とケース 6 の共通点は文書の中に二つの商品特性を含むことである。ケース 3 は「キャンメイク」(ブランド)と「落ち持ち」(評価表現)を含め、ケース 6 は「人気」(評価表現)と「オレンジ系」(色系)を含む。それに対し、他のケースはすべて一つの商品特性を述べた。

5.3 考察

表 3 で示したように、提案手法は検索タスク達成時間と達成クリック数を比較した結果、有意差があった。いわゆる、提案手法は検索する時間とクリック数の側面が比較手法 1 と 2 より検索効率が向上したことがわかる。それは、商品価値モデルの評価で示した高い適合率に関係がある。

比較手法 1 を利用する際、ユーザーの時間や手数を極めてかかることが見える。その要因を分析してみると、アットコスメに投稿するレビュー文が字数無制限により、500 字以上のレビューもたくさんある。だから、アットコスメはレビューを全文表示することではなく、「続き読む」ボタンが押されると全文を表示する。ユーザーは一目で多くのレビューを閲覧できるが、各レビューの一部しか読めない。ユーザーは欲しい情報を探索するために「続き読む」「戻る」ボタンをペアで繰り返し使うことがよくある。それに対し、Twitter に投稿するツイートは 140 字以内という制限があるため、一目で見通しやすい。それは効率向上の一つの理由として考えている。

また、アットコスメは美容総合サイトとして、アットコスメショッピングとアットコスメストアを広範で宣伝している。ユーザーはアットコスメにけるレビ

ユー検索する際、情報検索行為から商品販売行為まで移ることが期待されていることが予想できる。しかし、アットコスメは大量の商品宣伝広告を展示することにより、ユーザーの欲しい情報を見つけることが困難になってしまう。

比較手法2はTwitterキーワード検索である。周知のように、それはキーワードと一致したワードを含むツイートを検索できる。しかしながら、検索されたツイートの意味を考慮しないままユーザーに提供する。すると、図7のよう、一致したワードが含むが意味的にはユーザーに無効な情報が検索されたことがよくある。

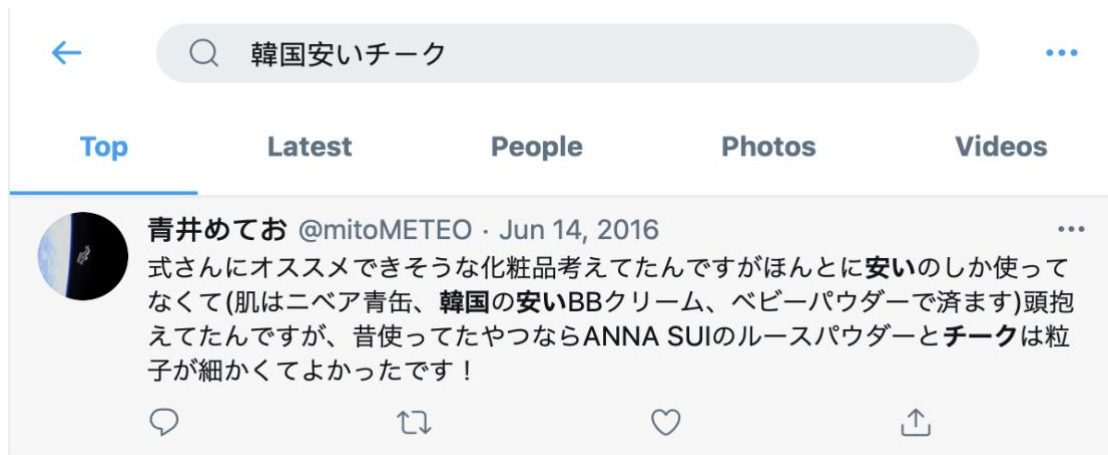


図 7 Twitter キーワード検索例

一般的に実験参加者はTwitterキーワード検索を利用する際、ユーザーがキーワードを調整しながら検索を繰り返す。なぜかという、キーワードの表現により「条件に該当するデータは存在しない」を提示することがあり、その場合には意味を維持することを前提としてキーワード表現を変更する。多くの実験参加者はキーワード表現を変更することに時間と手数をかけていた。

したがって、本研究では提案している検索手法は商品価値モデルによりレビューにタグをつけることにより、意味的に一致した情報を提供することで検索効率を向上できたと考えられる。

なお、ユーザーは予め設定した検索タスク（5.2.1 表7を参照）は主に二種類

に分類できる。（一）特定商品を入力し、商品特性が含まれているレビューを検索する。（二）商品特性を入力し、対応する特定商品が提示されているレビューを検索する。本実験の場合は、（一）に属しているのはタスク 2,5,6,8,15 であり、（二）に属しているのはタスク 1,3,4,7,9,11,12,13,14 である。これから、種別別に提案手法と比較手法 1（アットコスメ）の間に有意差があるかを検証する。検定結果は表 11 と表 12 のように示している。

	平均値(s)	標準偏差	平均値の差	t 値	p 値
提案手法	30.00	23.05	-49.00	-2.244	0.088
比較手法 1	79.00	36.47			

表 11 種類（一）達成時間比較

	平均値(s)	標準偏差	平均値の差	t 値	p 値
提案手法	7.78	4.92	-122.00	-3.458	0.009**
比較手法 1	129.78	108.98			

表 12 種類（二）達成時間比較

種類（一）は有意差がなかったことに対し、種類（二）は差異があることが明らかになった。

したがって、提案手法は種類（二）商品特性を入力し、対応する特定商品を提示するタイプのユーザーに効果的であるという結論が得られた。

本実験で提案手法がある程度で正確なスコアリング処理ができることを検証したが、問題点も明らかになった。

一つ目の問題点は商品価値モデルが商品レビューと商品広告を区別しないことである。ユーザーがレビュー検索システムを利用する際、商品に関連する詳細情報を了解することを重要な目的とするが、商品広告で商品情報を了解するとこの区別はユーザー実際の使用感想が読めることである。したがって、Twitter データを用いてレビュー検索手法はツイートの性質を分析する機能を加え、検索手法の精度を向上することが期待される。二つ目の問題点は商品価値モデルのトレーニング段階における、モデルの精度を追求するためにタグの総数を控えなければならないことである。例えば、「ブランド」というトピックに関する、ブランド名毎にタグをつけなかった。したがって、ブランド名を特定する検索の場

合、精確な検索結果が得られない状況となった。

5.4 結言

本章では提案手法の検索効率とレビュー推薦システムの正確性を検証するために行う評価実験の実験結果を挙げながら考察を述べた。実験結果は提案手法が比較手法より検索効率が高いことを示している。さらに、提案手法は商品特性を入力し、対応する特定商品が提示されているレビューを検索するユーザーに適用するという結論が得た。また、スコアリング処理が妥当であると言えるが、提案手法の問題点も明らかになった。今後、広告ツイートを分別すること及びブランドタグ数が増加することを検討する予定である。

第 6 章 おわりに

6.1 まとめ

近年、オンラインショッピングの普及とソーシャルメディア利用者の増加とともに、インターネット上の商品関連情報が膨大になっている。ユーザーが多くのレビュー情報を閲覧できるが、同時に欲しい情報を発見することは困難である。様々なレビュー検索公式サイトが広く使われており、複数のレビュー検索手法が提案しているが、まだ解決する必要がある問題点は存在している。Twitter は 140 字制限がある交流系・拡散系のソーシャルメディアであり、他のソーシャルメディアに比べると、「随意性」と「即時性」という特徴を持っている。大量のツイートに含まれるレビュー商品情報を効率的に利用することを目指し、本研究では Twitter データを用いて商品価値モデルによりレビュー検索手法を提案した。本研究の新規性はメイクアップ商品の特性によりレビュー文を区分けすることとユーザーが欲しい情報の含量によりレビューを自動的にスコアリング処理することである。

本研究では、ユーザーに豊富な商品関連情報を効率的に提供するため、ユーザー嗜好を考慮したメイクアップ商品を対象とする、Twitter データを用いた商品価値モデルによるレビュー検索手法を提案した。まず、研究対象（ファンデーション、リップスティック、チーク）をキーワードとする、ウェブクローラーを使用し Twitter データを収集しデータベースに保存した。次に、インターネットで掲載された既存の化粧品レビューから評価トピックと商品特性を抽出した。それから、学習用データ集を作成するため、人手で評価トピックと商品特性のタグを付けた。学習用データ集を用いてロジスティック回帰分析により商品価値

モデルを自動的に構築した。商品価値モデルの評価で示したのは、適合率について、「評価表現」は全て 90%未満であり、チークの「色系」は 88.94%であり、それ以外は全部 90%を超えた。一方、三つのカテゴリーの「季節」再現率は 63.68%から 86.91%であった。「季節」は 86.58%から 97.24%であった。「評価表現」についての評価結果は改善の余地が示した。特に、チークの「評価表現」の再現率は低い。その理由は、チークに関するレビューでは、評価表現として言及された単語が「発色」以外のものは滅多に見られない。

最後は商品価値モデルにより、Twitter データを対象とするレビュー検索システムを実現した。提案手法を評価するため、二つの評価実験を設計し行なった。その評価はレビューを効率的検索できているかと検索したレビュー順序の妥当性を検証した。

評価実験(1)は提案手法と比較手法を利用し、同様な検索タスクを達成する時間とクリック数を調べた。実験結果より提案手法のレビュー検索効率はよくなるという結果を得た。さらに、提案手法は商品特性を入力し、対応する特定商品が提示されているレビューを検索するユーザーに適しているという結果を得た。

評価実験(2)は MAP 評価指標を用いて提案手法で推薦するレビューの相関性を評価した。実験結果は提案手法の MAP 評価指標は 0.653 であった。その結果から、スコアリング処理は妥当であると言えるが、広告内容を取り除く機能が欠けていることやブランドのタグ総数が不足しているという問題点が判明した。

以上より、本研究で提案したレビュー検索手法はメイクアップ商品（ファンデーション、リップスティック、チーク）のレビュー検索する際、良い検索効率と正確なレビュー順番表示ができたことが明らかになった。特に、指定された商品

特性に対応する特定商品を検索するケースには効果的であることがわかった。

6.2 今後の展望

評価実験2の結果により、スコアリング処理を改善する必要があることが分かった。そのために、広告宣伝文と商品評価文を区別する課題と商品価値モデルの学習データ数を増加することを検討する予定である。

また、本研究では、Twitter データに含まれるレビュー情報を利用するために、ツイートを収集したが、テキストのみ収集した。しかしながら、Twitter には画像登録機能もあり、多くの利用者は画像を展示しながら投稿している。今後は画像情報を考慮したレビュー検索手法を検討することが課題である。

参考文献

- [1] 江崎大嗣, 川場真理子, 平野徹, レビュー文分類器を用いたレビュー文含有比率によるレビュー文判定, 2012-IFAT-105, No.2, (2012)
- [2] 乾孝司, 奥村学, テキストを対象とした評判情報の分析に関する研究動向, 自然言語処理, Vol.13, No.3, pp.201-241, (2006)
- [3] 相川直視, 山名早人, レビューからの商品比較表の自動生成, NLP2011, 発表論文集, Vol.17, pp.510-515, D2-3 Mar(2011)
- [4] アットコスメ: <https://www.cosme.net> (2021/02/01 アクセス)
- [5] 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, DEIM Forum B1-1, (2016)
- [6] 酒井美春, 松下光範, 上田真由美, 化粧品の評価項目別スコア生成のための評価表現辞書の自動構築, DEIM Forum B6-2s, (2019)
- [7] 白田由香利, 橋本隆子, 久保山哲二, インターネット上の口コミサイトにおける化粧品の評判分析, 学習院大学計算機センター年報, No.33, pp.2-7, Jul(2013)
- [8] 立石健二, インターネットからの評判情報検索, 自然言語処理, No.144-011, pp.75-82, (2001)
- [9] 二本木智洋, 住田一男, 文の構造化によるレビュー評価の分析・検索, 情報処理学会インタラクション, Vol.2002, No.7, pp.175-176, (2002)
- [10] 姚佳, 井戸田博樹, 原田章, インターネットのレビューか購買行動に及ぼす影響 -女子学生の化粧品購買のアンケート調査 から-, 経営情報学会 2014 年春季全国研究発表大会, pp.229-232, Jun(2014)
- [11] Duc-Hong Pham et al., “A least square based model for rating aspects and identifying important aspects on review text data,” The 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science, pp.265-270, (2015)

- [12] Zhang Bo, 白井清昭, レビューテキストの書き手の評価視点に対する評価点の推定, NLP2017 発表論文集, pp.803-806, Mar(2017)
- [13] 松波友稀, 上田真由美, 中島伸介, コスメアイテムの使用感および嗜好度判定を目的としたレビュー分析手法の提案, DEIM Forum 2015, pp.3-1, Mar(2015)
- [14] 駒田康孝, 山名早人, 商品評価ツイートからの属性語自動抽出手法の提案, DEIM Forum 2014, B5-6, (2014)
- [15] 杉浦広和, “議事録集合からの特徴語抽出とその応用に関する研究”, DEIM Forum C2-3, (2016)
- [16] 大塚裕子, 自由記述アンケート回答の意図抽出および自動分類に関する研究, 神戸大学大学院自然科学研究科博士論文, Mar(2004)
- [17] 中山記男, 江口浩二, 神門典子, 感情表現の抽出手法に関する提案, 情報処理学会研究報告, 自然言語処理, Vol. 2004, No.108, pp.13-18, Nov(2004)
- [18] 立石健二, インターネットからの評判情報検索, 自然言語処理, No.144-011, pp.75-82, (2001)
- [19] 磯沼大, 藤野暢, 浮田純平, 村上遥, 浅谷公威, 森純一郎, 坂田一郎, 文書分類とのマルチタスク学習による重要文抽出, JSAI2017, May(2007)
- [20] 橋本泰一, 村上浩司, 乾孝司, 内海和夫 石川正道, 文書クラスタリングによるトピック抽出及び課題発見, 社会技術研究論文集, Vol.5, 216-226, Mar. (2008)
- [21] 赤羽根隆宏, 野沢理恵, 小杉涼夏, 木村紗耶香, 佐藤央, 荒井正之, 海上隆, 品詞構造と tf-idf 法に基づく感性語の自動選定法の提案, 日本感性工学論文誌, Vol.8, No.3, pp.691-699, (2009)
- [22] Stefan Büttcher, Charles L. A. Clarke, Gordon V. Cormack, 『情報検索: 検索エンジンの実装と評価』 森北出版 (2020)
- [23] 松河秀哉, 大山牧子, 根岸千悠, 新居佳子, 岩崎千晶,, 堀田博史

- トピックモデルを用いた授業評価アンケートの自由記述の分析, 日本教育工学会論文誌, Vol41, No.3, pp.233-244,(2018)
- [24] 村松亮介, 横山昌平, 福田直樹, 石川博, 検索クエリ分類モデルに基づく多視点クラスタリング検索エンジンの評価, DEIM Forum B9-4, (2010)
- [25] 奥中大地, 徳丸正孝, ニュートラルネットワークを用いた感性検索モデル, 日本感性工学論文誌, Vol.11, No.2, pp.331-338,(2012)