

商品の利用目的の特徴表現と検索への応用

村本 直樹[†] 湯本 高行[‡] 大島 裕明[†]
Naoki Muramoto Takayuki Yumoto Hiroaki Ohshima

1. はじめに

様々な Web サービスが提供されるようになり、多くのユーザによって EC サイトや、スポット検索・予約サービスなどが利用されている。同時に、利用ユーザにより、商品レビューをはじめ、飲食店のレビューや観光地のレビューなど様々なテキストデータが投稿され、これらのレビューを対象とした分析が盛んに行われている。本研究で私達は 図 1 のような商品レビューを分析対象とし、レビュー内の「商品の利用目的」に注目する。例えば、「シューズ」について詳しくないユーザが、「神戸マラソンに初めて出るのでシューズがほしい」と具体的な欲求を持ち、EC サイト上で単純に「神戸マラソン 初めて」と検索しても、適した商品にはヒットしないといった問題がある。また、商品の利用目的に注目した代替性についても考える。例えば、商品「シューズ」において「ダイエット」と、「筋トレ」という利用目的は比較的近いものであるが、それ意識した検索を行うことはできない。他にも、「バスケットボール」をするための「バスケットシューズ」が、靴のクッションが良い点や、足首のサポートが強いことなどから、「ウェイトトレーニング」にも利用できると言われている。このように、商品の利用者が、商品の製造者の意図とはすこし異なる目的でその商品を利用していることもある。

ここで、本研究では商品レビューや商品名から、「利用目的」に注目して商品をベクトルで表現する。そしてそのベクトルを基に、商品検索モデルを作成する。

しかし、本研究で扱う商品レビューには以下の要素が、

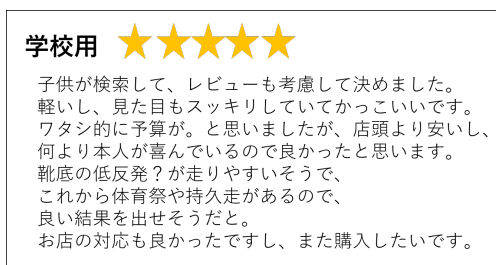


図 1: シューズに関する商品レビュー例

混在していると考えられる。

- 商品の利用目的（本研究の主な対象）
- 商品に関する評価（対象となりえる）
- 商品購入に至る背景
- 販売者側の対応や配送に関する評価

このように、テキストデータを扱う場合、商品レビューに関わらず、取り組む問題に対して必要となるのは、テキストの一部であることが考えられる。例えば、図 1 のレビューにおいては、「体育祭」や「持久走」という利用目的が存在している。また、これらの利用目的において、「靴底」や「低反発」、「走りやすい」といった、商品の評価に関する表現も重要である可能性がある。これは、ある商品のある利用目的で使用する際に、必要となる商品の注目観点を表していると考えられる。つまり本研究において重要になる語は以下のとおりである。

- 利用目的そのものを表現する語
- ある商品を何らかの利用目的で使用する際に必要となる商品の性能（商品の注目観点）を表現する語

逆に図 1 において、以下の文は「商品の利用目的」に関しては、意味のない内容であると考えられる。

- 子供が検索して、レビューも考慮して決めました。
- ワタシ的に予算が。と思いましたが、店頭より安いし、何より本人が喜んでいるので良かったと思います。
- お店の対応も良かったですし、また購入したいです。

このような問題がテキストデータを扱う際に存在しており、Baf-of-Words や TF-IDF, Word2Vec[1] などのベクトル化手法を単純に用いてしまうと、本来の注目観点とは違う点についての表現も、同様に扱われてしまう問題点がある。

重要なテキストを抽出するために、我々がこれまでに行ってきた研究の 1 つは、ソーシャルメディア上から構文パターンを用いて、商品の利用目的を発見するものである [2]。ここで利用している構文パターンは、クエリに対する同位語を、前後に出現する構文パターンを組み合わせる事により、発見した手法 [3] を応用したものである。

本研究ではまず、「商品の利用目的を特徴づける表現（特徴表現）」を上記の両方向構文パターンを用いて抽出

[†] 兵庫県立大学 応用情報科学研究科, Graduate School of Applied Informatics, University of Hyogo

[‡] 兵庫県立大学 工学研究科, Graduate School of Engineering, University of Hyogo

する。続いて、各商品ジャンル毎の Word2Vec モデルを用いて特徴表現の拡張を行う。そして、商品に投稿されたレビューを特徴表現のみを用いてベクトル化する。そのベクトルを基に「利用目的」による商品検索に応用する手法を提案する。

以降、2章では関連研究について、3章では特徴表現の説明、対象とするデータと前処理について、4章で商品の利用目的を特徴づける表現の取得手法と取得結果について、5章で特徴表現を基に各商品をベクトルで表現する手法と検索結果について、6章でまとめと今後の課題について述べる。

2. 関連研究

本研究では、商品レビューや商品名のテキストデータを基に、「利用目的」を特徴づける表現を取得し、商品ジャンル毎の特徴表現集合を作成する。そして、その特徴表現を用いて、商品をベクトルで表現することで、利用目的による商品検索を行うことを目標としている。

一般的に、EC サイトでは協調型の商品推薦システムを利用している事が多い [4][5]。そのため、利用目的による検索は難しい。

ユーザの目的という観点について研究されているものに、高見らの研究 [6] が挙げられる。高見らは Web 検索結果において、検索クエリからユーザの検索目的にあったスニペットについて分類をするに取り組んだ。また、徳田ら [7] は、観光地検索において、明確な目的がないユーザに対し、発見思考型のナビゲーションシステムを提案している。

利用目的に注目し、商品の推薦システムを構築する研究も行われている。本田ら [8] は、EC サイトにおける商品カテゴリに注目して、カテゴリ特徴ベクトル、商品特徴ベクトル、レビュー特徴ベクトルを作成し、類似度を算出することで、ユーザの使用目的にあった代替商品の推薦システムを提案した。McAuley ら [9] は、ある商品を購入する際に推薦すべき商品について、その商品について代替利用できる代替品と、その商品と同時に買うことが推薦される補完品の2種類に注目し、代替品について推薦システムを提案している。

利用目的とはその商品を使った人の行動と言えるが、テキストコーパスから人々の行動を取得する研究も行われている [10] [11] [12]。たとえば、馬縹ら [12] は、「薬剤師が薬を調合する」のような、ある職業における行動を、対象の職業が主語になっている主語ベースの文と、対象の職業に従事するユーザによって書かれた著者ベースの文から取得している。

本研究では Q&A サイトのテキストコーパスを利用しているが、Q&A サイトの分析を行う研究も行われてい

る [13]。

また、ユーザの意思決定に注目し、商品の推薦システムを構築した研究も存在する。小池ら [14] は、実際に買い物をする際の女性と男性の差に注目し、キーワードを可視化した対話型推薦アルゴリズムを提案した。益田ら [15] は、実際の家電量販店の販売員の商品販売過程を、「販売員思考」、「対象ユーザ」、「対象商品」「システムの特徴とコンセプト」の4段階で構成しモデル化することで商品推薦システムを考案している。

3. 商品レビューと特徴表現について

3.1 商品の利用目的を特徴づける表現

1章で述べたように、本研究では商品の利用目的を特徴づける表現を取得し、各商品ジャンル毎に、特徴表現集合を作る。この「商品の利用目的を特徴づける表現」の具体例を述べる。以下のように、同一のシューズに対する2件のレビューが存在したとする。

「とてもお店の対応は早かったです。私は特にソールが分厚いので気に入っています。週末に軽くランニングしているので利用しています。」

「店の対応には満足しています。ソールが特徴的で、足に優しいので普段使いにしています。」

このとき1つ目のレビューにおいて、利用目的を直接的に表す表現は「ランニング」である。さらに、「週末」や「軽い」といった、より詳しい状況を表す表現も含まれている。また、「ソール」、「分厚い」といった商品情報についての表現も、「シューズ」を「ランニング」で利用するために特徴づけていると考える。2つ目のレビューにおいては、「普段使い」という利用目的と、「ソール」、「優しい」という、利用目的を表現する商品情報が含まれている。これらの表現を図2のように、それぞれのレビューからまとめて特徴表現として取得する。特徴表現を取得するための手法を次章以降で述べるが、主なステップは以下の2つである。

1. 両方向構文パターンによるシードとなる特徴表現の抽出
2. Word2Vec モデルを用いた特徴表現の拡張
3. 各種ワードセットによる特徴表現の追加

まずはじめに、構文パターンを用いた抽出手法により、「ランニング」や「ジム」、「ダイエット」などの直接的に利用目的を表現する語を抽出する。ここで取得された特徴表現を、「シード特徴表現」とする。その後、各商品ジャンル毎の Word2Vec モデルと「シード特徴表現」より、特徴表現を拡張することで、各商品ジャンル毎に

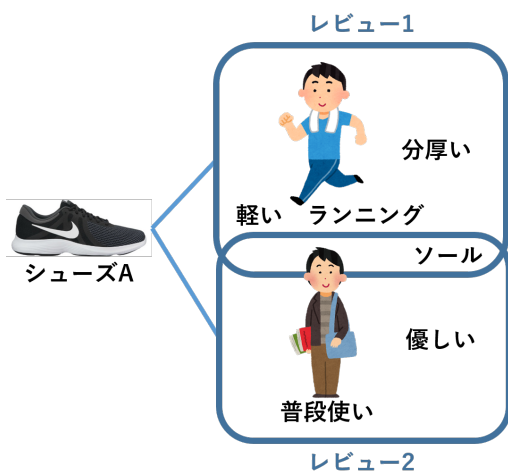


図 2: 商品と利用目的を表現する特徴表現

表 1: 楽天市場 商品レビューの各データ量

商品	レビュー数	商品 ID 数
シューズ	699,843	96,713
シャツ	1,174,437	257,053
リュック	238,464	27,362

特徴表現集合を作成する。ワードセットは 2 種類作成する。1 つ目は商品レビューに頻出する「配送」や「購入」に関する一般語を、全ジャンルの商品レビューを用いた Word2Vec モデルより作成する。2 つ目は、各商品ジャンル毎に商品の属性や観点を表す語を、各商品ジャンルのレビューデータを用いて作成する。

3.2 対象とした商品レビューデータ

本研究では、商品レビューと商品名を基に、商品の利用目的を特徴づける表現を取得し、各商品ジャンル毎の特徴表現集合を作成する。そして、その特徴表現を用いて商品をベクトルで表現する。商品レビューのコーパスは「楽天データ公開」*にて公開されている「楽天市場」[16]に投稿された商品レビューを利用する。本手法では、前節で述べたように、各商品ジャンル毎の Word2Vec モデルと、一般語のワードセットための Word2Vec モデル用のコーパスを分けて利用する。まず、各商品ジャンル毎のコーパスについて述べる。利用する商品ジャンルは「シューズ」、「シャツ」、「リュック」の 3 種類であり、2010 年 1 月から 2012 年 12 月の間に投稿されたレビューである。それぞれのレビュー件数と、どの商品について書かれたものであるかを判断する商品 ID の件数を表 1 に示す。

次に、「配達」に関するワードセットや、「購入」に関する

ワードセットなどを作成するためのコーパスに、2012 年の 1 年間分の商品レビューすべてを使用する。1 年分の商品レビューの件数は、26,042,154 件である。

3.3 文書の前処理

本手法では取得する特徴表現は 1 つの形態素からなるものとする。そのためにもまず与えられた文書を以下の手順で前処理を行う。

1. 「。」「!」「?」以外の記号の削除
2. 英数文字の半角化
3. 数字を 0 に統一
4. 「。」「!」「?」で文単位に分解する。
5. 配送に関するワードセットの語を含む文の削除
6. 形態素単位に分ち書き

分ち書きは表層系で形態素単位に分ち書きを行う。形態素解析器には MeCab[†]を用いる。MeCab の辞書については新語に対応している mecab-ipadic-NEologd[‡] (2019 年 1 月 31 日更新分) を追加している。この前処理を行ったレビューデータを用いて構文パターンによる特徴表現抽出、Word2Vec モデルの作成を行う。

4. 特徴表現の取得

本研究では主に以下の 2 点の問題に取り組んでいる。

- ある商品における特徴表現の取得
- 得られた特徴表現を用いて商品をベクトルで表現

まず、1 点目の問題について本章で述べ、2 点目の問題については次章で述べる。

4.1 両方向構文パターン

本節で、本研究で利用する両方向の構文パターンについて述べ、次節で両方向の構文パターンによる、シード特徴表現の抽出手法について述べる。

まず、商品レビュー上のテキストにおいて、利用目的を直接的に表す語の、前後に出現する構文パターンを見つける。抽出したい特徴表現の前に出現する形態素列からなる構文パターンを「前方パターン」。同様に特徴表現の後に出現する形態素列からなる構文パターンを「後方パターン」とする。

- <前方パターン> target
- target <後方パターン>

後方パターンにおける具体例は、「ランニング用」や「トレーニング時に」といった、特徴表現から「用」「時に」につながるものである。

*https://rit.rakuten.co.jp/data_release_ja/

[†]<http://taku910.github.io/mecab/>

[‡]<https://github.com/neologd/mecab-ipadic-neologd>

構文パターン	例
の(ため 場合 ばあい)	ランニングのため
用 向け むけ	ランニング用
の?(時 とき)に?	ランニングの時
[でに]使[あ-ん用]	ランニングで使う
[でに][活利]用	ランニングで利用

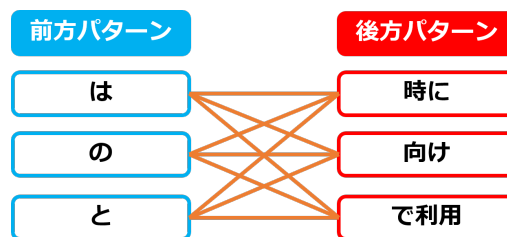


図 4: 構文パターンの組み合わせ例

構文パターン	例
の	～の体育会
と	～とランニング
は	～はサッカー
、(読点)	、散歩
^※正規表現	^ランニング

表 4: Word2Vec のパラメータ

パラメータ	設定
vector	300
window	5
min_count	5
model	Skip-gram

前方パターンにおいては、「～はランニング」や、「～の体育会」のような助詞が多く見られた。また、文頭や読点から、「ランニングのときに履いています!」のように始まるものも多く存在していたため、前方パターンに追加した。これらの両方向の構文パターンについて、利用するパターンを表 2、表 3 に示す。

4.2 両方向構文パターンを用いた挟み込み抽出

本節では、両方向の構文パターンを用いた特徴表現の抽出手法について述べる。手法の基本となるのは我々が行ってきた手法のうち、両方向の構文パターンによる挟み込み抽出手法である。本研究では抽出する特徴表現は、1つの形態素からなるものとする。前方パターンと後方パターンを、図 3 のように、特徴表現を挟み込むように組み合わせる。この時、両方向の構文パターンの組み合わせ方は、図 4 のように掛け合わせて用いる。この時抽出された特徴表現から、ストップワードとひらがなのみのものを削除し、出現頻度上位 1,000 件をシード特徴表現とする。ストップワードに関しては、4.4 節で述べる。



図 3: 両方向構文パターンによる挟み込み抽出

4.3 Word2Vec を用いた特徴表現の拡張

本節では、得られたシード特徴表現から、各商品ジャンル毎の Word2Vec モデルを用いて拡張する手法について述べる。また、実際に拡張される語について、全ジャンルで作成する Word2Vec モデルを用いた場合との比較を行う。

「筋トレ」や「筋肉トレーニング」など、同じ意味でも表記ゆれが起こる特徴表現は多く存在する。このような特徴表現は、それぞれのコーパス内での出現頻度が低くなり、構文パターンにより抽出できない可能性がある。また、具体度の高い「カジュアルフライデー」なども、出現頻度が低いことから抽出が困難になる。そのため、シードとなる特徴表現と似た表現、また、その商品ジャンルにおいて近い意味をもつ特徴表現を取得するために、以下の手順で特徴表現の拡張を行う。

1. 各商品ジャンル毎に Word2Vec モデルの作成
2. シードとなる特徴表現とのコサイン類似度トップ 10 で値が 0.5 以上の語を追加
3. 天気と商品ジャンル毎のワードセットを追加

まずはじめに、3 章で述べた各商品ジャンル毎のデータセットを用いて、表 4 のパラメータで Word2Vec モデルを作成する。

この Word2Vec モデルを用いて、シードとなる特徴表現 1,000 件に対し、コサイン類似度が高い 10 件のうち、類似度が 0.5 以上のものを特徴表現に追加する。そして最後に、天気や各商品ジャンル毎に予め用意したワードセットを追加し各商品ジャンルの特徴表現集合とする。この手法で拡張することにより得ることができた特徴表現を表 5 にまとめる。ワードセットについては次節で述べる。

表 5: 拡張を行うことで得られた特徴表現の例

シューズ	シャツ	バッグ
オールウェザー おしゃれ	ウェイトトレーニング カジュアルフライデー	高尾山 カジュアルデイ
神戸マラソン ウエッジソール	ガリガリ ドレスシャツ	ガールスカウト 耐水
タップダンス	ヘビーユーズ	行楽シーズン

ここで、同じパラメータで、2012年1年分の商品レビューすべてを用いて作成されるの Word2Vec モデルと、各商品ジャンル毎の Word2Vec により得られる拡張される特徴表現の比較を行う。

商品ジャンル「シューズ」に対して、「マラソン」、「ダイエット」、「土」の3つの語について、それぞれのモデルでコサイン類似度の上位5件を取得する。結果を表6に示す。表6を見ると、「マラソン」について、全ジャンルを用いた場合は、ECサイトのセール名の名前として扱われ、「駅伝」以外は、関係のないものである。しかし、シューズのレビューのみを用いたものは、正しく「走る」マラソンに関係する特徴表現になっており、「神戸マラソン」のような具体的な例も含まれる。次に、「ダイエット」について、全ジャンルを用いた場合は、「ファスティング」や「食事制限」などの様々な「ダイエット」に関する語が存在している。シューズのレビューのみをもちいたものは、「シェイプアップ」や「エクササイズ」、「筋トレ」など、シューズを履いた上で「ダイエット」するための特徴的な表現が取得できている。最後に「土」について見ると、全ジャンルを用いた場合は、「土」の種類に関する語が取得されている。これは、「シューズ」という観点からはあまり意味のないものである。しかし、シューズのレビューのみをもちいたものは、「シューズ」を履いて走る地面の状況を表す特徴表現が取得できている。

このように、各商品ジャンル毎に Word2Vec モデルを作ることにより、その商品ジャンルにおける特徴表現の類似性を図ることが可能になり、シード特徴表現より拡張することができた。

4.4 ワードセットの作成

本節では、ストップワードや、特徴表現に追加するためのワードセットと、作成方法について述べる。作成するワードセットは大きく分けて以下の2種類である。

1. 一般的な語のワードセット
2. 各商品ジャンル毎の主な属性を表すワードセット

1つ目の一般的な語のワードセットとは、商品レビュー全体に頻出する「配送」や「購入」、「店」、「天気」に関連する語の集合である。これらを3.2節で述べた、1年

間のすべてのレビューを用いた Word2Vec モデルより作成する。作成する際の手順を以下に示す。

1. シードとなる語を選択する。
2. シード語に対してコサイン類似度の高い上位100件を取得する

また、各商品ジャンル毎にもワードセットを作成する。このワードセットは各商品ジャンル毎の Word2Vec モデルを用いて作成する。商品ジャンル毎にいくつかの属性を表す語をシード語として与え、上記の方法と同様に類似度が高い50件を取得する。これらのシードとなる語を表7に示す。

作成したワードセットの内、ストップワードとして扱うもの、特徴表現に追加するワードセットとして扱うものの2つに分ける。詳細を以下に示す。

- ストップワード：配送・店、購入、色
- 追加するワードセット：天気、各商品ジャンル毎の属性ワードセット

5. 商品検索への応用

5.1 ベクトル化手法と商品検索モデルの作成

本節では、4章で得られた特徴表現集合を用いて、商品をベクトル化する手法と、商品検索への応用方法について述べる。いくつかのクエリに関して実際に得られる結果と、考察については次節で述べる。

まずはじめに、商品レビューデータ中の商品IDを基に、商品IDに対して、商品名と複数存在するレビューデータをマージしたテキストを紐付ける。そして、マージしたテキストを分かち書きして、各商品ジャンル毎の特徴表現集合に含まれる語のみを残す。そしてTF-IDFを用いてモデルを作成し、各商品をベクトルで表現する。手順を以下にまとめる。

1. 商品名と複数の商品レビューをまとめる
2. まとめたテキストを分かち書き
3. 特徴表現のみを残す
4. TF-IDFを用いてモデルを作成しベクトル化

表 6: シューズと全ジャンルを用いたときの Word2Vec に類似度の高い語の違い

	マラソン		ダイエット		土	
コーパス	全ジャンル	シューズ	全ジャンル	シューズ	全ジャンル	シューズ
1位	楽天スーパーセール	駅伝	減量	シェイプアップ	用土	グラウンド
2位	駅伝	フルマラソン	プチダイエット	エクササイズ	培養土	グラウンド
3位	ポイントアップ	ウルトラマラソン	痩せる	筋トレ	堆肥	芝
4位	タイムセール	ランニング	ファスティング	引き締め	腐葉土	砂
5位	セール	神戸マラソン	食事制限	健康	赤玉土	芝生

表 7: ワードセット作成のためのシードとなる語

	シード語			
配送・店	配送	発送	配達	
購入	購入	買う	品切れ	
色	赤	レッド	あか	
天気	晴れ	雨	曇り	
シューズ	クッション	アッパー	ソール	重さ
シャツ	素材	通気性	大きさ	丈
バッグ	容量	デザイン	大きさ	重さ

続いて、自由記述のクエリによる検索手法とその結果について述べる。自由記述で書かれたテキストクエリに対して、上記のベクトル化手法と同様に、分かれ書きを行い、特徴表現のみを残す。そして、作成済みの TF-IDF モデルを利用して、特徴表現のみで表現されるクエリをベクトル化する。最後に、得られたクエリベクトルと、商品 ID 毎のベクトルとのコサイン類似度が高いものから、検索結果を表示する。

5.2 検索結果と考察

実際に商品「シューズ」に対して、「神戸マラソンで初めて走る」というクエリで検索した場合について、「楽天市場」で検索した結果と作成した検索モデルで検索した結果を比較する。「楽天市場」の検索機能を利用して、「スポーツ・アウトドア」のジャンルでフィルターを掛け、「神戸マラソンで初めて走る」というクエリで検索を行う。検索結果を見ると、どのアイテムにもヒットしない。クエリを「神戸マラソン 初めて」や、「神戸マラソン」のみにしても、ランニングシューズにはヒットしない。

続いて、作成した検索モデルを用いて同様のクエリで検索を行う。まず、「神戸マラソンで初めて走る」というクエリに対し、「神戸マラソン」、「初めて」、「走る」の3つが特徴表現として選択され、TF-IDF モデルよりクエリがベクトル化される。このクエリベクトルを用いて、コサイン類似度の高い商品を検索した結果の上位5件を表8に示す。表8を見ると、各スポーツメーカーのランニングシューズが検索できていることがわかる。詳細

について調べると、3位の「PUMA ウィメンズ コンプリートフウジン」のみ、中上級者向けのシューズであったが、その他はエントリーモデルのランニングシューズであることが確認でき、マラソンで初めて走るという利用目的を、満たすことができるシューズである可能性が高いと考えられる。

しかし、「野球 初心者 スパイク」のようなクエリを入力した際に、「楽天市場」では適した商品がヒットするが、作成した検索モデルでは表9のように、5位の「ZETT 野球スパイク」以外、クエリに適した結果が検索できていない。これは、商品名に「【野球__sale】」という、セールに関する情報が含まれていることによる間違いである。本手法では、商品名も含めた特徴表現から TF-IDF により、特徴ベクトルを作成しているため、商品名の中に「野球」が含まれている上記のような場合、近い商品としてしまい、不適切な結果になってしまったと考えられる。

6. まとめと今後の課題

本研究では商品レビューから、その商品の利用目的を特徴づける表現の取得を、両方向構文パターンと各商品ジャンル毎の Word2Vec モデルを用いることで実現した。各商品ジャンル毎に Word2Vec モデルを作成することで、その商品ジャンルにおける利用目的として類似度の高い、特徴的な表現を取得することができた。そして得られた特徴表現より商品をベクトルで表現し、利用目的による商品検索を行うことを実現した。

今後の課題は、まず取得された特徴表現を用いて商品をベクトル化する際に、1つの商品に対して複数存在するレビューの扱い方について検討する。今回は商品名と全レビューを一つにまとめて扱ったが、それぞれでベクトル化を行うことを試みる。また、最終的なベクトル化手法として TF-IDF を用いたが、各商品ジャンル毎の Word2Vec モデルを再度利用することも検討する。最語に評価についても指標を用いた定量的評価や、ユーザ評価を行っていく予定である。

表 8: 「神戸マラソンで初めて走る」の検索結果

順位	商品名 (一部)
1位	ナイキ ブームストラクチャー+15 シールド ※ランニングシューズ
2位	アシックス フェザーファイン ワイド ※ランニングシューズ
3位	PUMA ウィメンズ コンプリートフウジン ※ランニングシューズ
4位	アシックス GT-2170 NEW YORK-SW ※ランニングシューズ
5位	アシックス ゲルカヤノ ※ランニングシューズ

表 9: 「野球 初心者 スパイク」の検索結果

順位	商品名 (一部)
1位	ナイキ トータル 90 レーザー 3 ※サッカースパイク
2位	ANNA COLLECTION クロスベルトウェッジヒールパンプス
3位	リラックマ スリッパ
4位	ASBees カジュアルシューズ
5位	ZETT 野球スパイク

謝辞

本研究の一部は JSPS 科学研究費助成事業 JP16H02906, JP18H03494, JP17H00762, JP16H01756, JP18H03243 による助成を受けたものです。また、本研究では国立情報学研究所の IDR データセット提供サービスにより楽天株式会社から提供を受けた「楽天データセット」を利用しました。ここに記して謝意を表します。

参考文献

- [1] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” *Advances in neural information processing systems*, pp.3111–3119, 2013.
- [2] 村本直樹, 湯本高行, 大島裕明, “ソーシャルメディアからの商品の利用目的の発見,” 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2019), pp.B4–5, 2019.
- [3] 大島裕明, 田中克己, “両方向構文パターンを用いた web 検索エンジンからの高速関連語発見手法,” *日本データベース学会論文誌*, vol.7, no.3, pp.1–6, 2008.
- [4] G. Linden, B. Smith, and J. York, “Amazon. com recommendations: Item-to-item collaborative filtering,” *IEEE Internet computing*, vol.7, no.1, pp.76–80, 2003.
- [5] 麻生英樹, 小野智弘, 本村陽一, 黒川茂莉, 櫻井彰人, “協調フィルタリングと属性ベースフィルタリングの統合について,” *電子情報通信学会技術研究報告. NC, ニューロコンピューティング*, vol.106, pp.55–59, 2006.
- [6] 高見真也, 田中克己, “検索目的に基づくスニペットの動的再生成によるウェブ検索結果の個人適応化,” *日本データベース学会論文誌 DBSJ Letters*, vol.6, no.2, pp.33–36, 2007.
- [7] 徳田英隼, 伊藤昌毅, 高汐一紀, 徳田英幸, “ぶらりナビ: 潜在的欲求を引き出す発見志向型ナビゲーションシステムの構築,” *マルチメディア, 分散, 協調とモバイル (DICOMO2006) シンポジウム*, pp.485–488, 2006.
- [8] 本田達也, 北山大輔, 角谷和俊, “オンラインショッピングサイトにおけるレビューを用いた商品対の目的判定による商品推薦,” *平成 24 年度 情報処理学会関西支部支部大会*, pp.G–28, 2012.
- [9] J. McAuley, R. Pandey, and J. Leskovec, “Inferring networks of substitutable and complementary products,” *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785–794, 2015.
- [10] E. Filatova and J. Prager, “Tell Me What You Do and I’ll Tell You What You Are: Learning Occupation-related Activities for Biographies,” *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp.113–120, 2005.
- [11] Z. Kozareva, “Learning verbs on the fly,” *Proceedings of COLING 2012: Posters*, pp.599–610, 2012.
- [12] 馬縹美穂, 笹野遼平, 高村大也, 奥村学, “職業ごとの行動に関する知識の収集,” *情報処理学会論文誌データベース (TOD)*, vol.11, no.3, pp.12–22, 2018.
- [13] N. Aikawa, T. Sakai, and H. Yamana, “Community qa question classification: Is the asker looking for subjective answers or not?,” *情報処理学会論文誌データベース (TOD)*, vol.4, no.2, pp.1–9, 2011.
- [14] 小池恵里子, 伊藤貴之, 渡辺知恵美, “提示量を適正化した女性向け商品検索支援システム,” 第 5 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2013), pp.D2–1, 2013.
- [15] 益田怜央, 増田英孝, 山田剛一, 福原知宏, “家電製品の利用シーン, 要求機能, 製品体系の知識に基づく商品推薦システム: 実装と評価,” *情報処理学会情報基礎とアクセス技術 (IFAT) 研究会*, vol.2014-DD-93, no.6, pp.1–8, 2014.
- [16] 楽天株式会社, “楽天データセット. 国立情報学研究所情報学研究データリポジトリ. (データセット),” 2014. <https://doi.org/10.32130/idr.2.0>.