

# 利用目的による商品検索のための 楽天レビューからの特徴表現抽出

村本直樹, 湯本高行, 大島裕明 (兵庫県立大学)

## 神戸マラソンに初めて出るから シューズがほしい!

### 背景

既存のECサイトでは

シューズ

神戸マラソンで初めて走る

検索結果0件

お探しの商品は見つかりません

利用目的を入力して  
検索できたらいいなあ



### 提案手法

アイテム毎の商品レビュー  
(楽天市場)

週末にジムでランニングするのに利用しています。  
購入前は底が厚いのでどうかなくと思いましたが、  
履いてみたらクッション性にびっくり!  
最近には父にも薦めています。  
お値段はそこそこしますが買い得だと思います。

利用目的を特徴づける表現  
(特徴表現)のみを残す

各アイテムのベクトル化

### 商品ジャンル毎に特徴表現集合を作成

| 前方パターン | 特徴表現   | 後方パターン |
|--------|--------|--------|
| は      | ランニング  | 時に     |
| は      | トレーニング | 時に     |
| は      | 散歩     | 時に     |

前後の構文パターンの挟み込みより  
シードとなる特徴表現の取得

上位1,000件をシード語として取得  
ひらがなだけの語、  
予め用意したストップワードを削除

前方 後方  
は 時に  
の 向け  
と で利用  
構文パターンの  
組み合わせ方の例



各商品ジャンル毎に  
Word2Vecモデルより  
特徴表現の拡張

取得されたシード特徴表現を基に、  
コサイン類似度がTOP10かつ0.5以上の語を  
特徴表現に追加する。

アイテム毎のレビュー集合と商品名から  
特徴表現のみを残す

TF-IDFでアイテムのベクトル化  
(モデルの作成)

クエリから特徴表現のみを残し  
作成済みモデルよりベクトル化

コサイン類似度より検索

### 特徴表現の例 (シューズ)

| シード語                                                                       | +α (拡張)                                                          |
|----------------------------------------------------------------------------|------------------------------------------------------------------|
| ランニング<br>ダイエット<br>ジョギング<br>ウォーキング<br>マラソン<br>ガーデニング<br>フットサル<br>歩く<br>送り迎え | 季節<br>天気<br>朝<br>夜<br>山<br>六甲山<br>雪国<br>海<br>体育祭<br>音楽会<br>野外フェス |

| 構文パターン<br>(正規表現) |             |
|------------------|-------------|
| 前方               | 後方          |
| の                | のため 場合 ば あい |
| と                | 用 向け むけ     |
| は                | の? 時 とき に?  |
| 、(読点)            | [で]に 使 あん 用 |
| ^ (文頭)           | [で]に [活利] 用 |

ジャンル毎のコーパスを  
用いることで適した  
特徴表現を取得

| Word2Vecモデルのコーパスによる拡張語の違い |          |         |         |     |       |
|---------------------------|----------|---------|---------|-----|-------|
| マラソン                      |          | ダイエット   |         | 土   |       |
| 全体                        | シューズ     | 全体      | シューズ    | 全体  | シューズ  |
| 楽天<br>スーパーセル              | 駅伝       | 減量      | シェイプアップ | 用土  | グラウンド |
| 駅伝                        | フルマラソン   | プチダイエット | エクササイズ  | 培養土 | グランド  |
| ポイントアップ                   | ウルトラマラソン | 痩せる     | 筋力      | 堆肥  | 芝     |
| 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     |
| セール                       | 神戸マラソン   | 食事制限    | 健康      | 赤玉土 | 芝生    |

拡張することで表記ゆれや  
間接的に利用目的を  
特徴づける表現を取得

| 拡張により得られた特徴表現の例 |            |         |
|-----------------|------------|---------|
| シューズ            | シャツ        | リュック    |
| オールウェザー         | ウェイトトレーニング | 高野山     |
| おしゃれ            | カジュアルフライデー | カジュアルディ |
| 神戸マラソン          | ガリガリ       | ガールスカウト |
| ウェッジソール         | ドレスシャツ     | 耐水      |
| クラブダンス          | ヘビークース     | 行楽シーズン  |

### 結果

シューズ

神戸マラソンで初めて走る

| 順位 | 商品名                      |
|----|--------------------------|
| 1位 | ナイキ ズームストラクチャー +15シールド   |
| 2位 | アシックス フェザーファイン ワイド       |
| 3位 | PUMA ウィメンズコンプリートフウジン     |
| 4位 | アシックスGT-2170 NEW YORK-SW |
| 5位 | アシックスゲルカヤノ               |



商品名に「野球セール」が  
含まれていたため、野球の  
スパイクではない商品がヒット

シューズ

野球の試合に初めて出る

| 順位 | 商品名                          |
|----|------------------------------|
| 1位 | ANNA COLLECTION ヒールパンパス      |
| 2位 | アグラ【AGLA】BANDIERA ※フットサルシューズ |
| 3位 | シースリーJF ※サッカーシューズ            |
| 4位 | リラクマスリッパ                     |
| 5位 | ASBees カジュアルシューズ             |

### 今後の課題

アイテムのベクトル化手法について再検討する。現状、単純にTF-IDFを用いているため検索においては「神戸マラソン」と「マラソン」の類似度を利用できていない。また「野球スパイク」の例のように1語に引っ張られてしまう事も起きている。この問題を解決するために、作成済みのジャンル毎のWord2Vecモデルを活用することを検討している。また定量評価も行う予定である。

エンタリーモデルの適した  
シューズ(3位以外)が  
検索できた