

異なる観点を考慮した商品説明に対応した商品レビューの抽出

佐藤 風希[†] 鈴木 優[†]

[†] 岐阜大学工学部電気電子・情報工学科 〒501-1193 岐阜県岐阜市柳戸1番1

E-mail: †y3033072@edu.gifu-u.ac.jp, †ysuzuki@gifu-u.ac.jp

あらまし 本研究は商品説明に対応した商品レビューを冗長性を減らして表示することを目的とする。商品説明や商品レビューには、異なる観点の商品説明にも用いられる単語を含んでおり、単語だけでは商品説明ごとに分けられない場合がある。このような場合においても、文脈によって同じ単語でも異なる分散表現を得られるBERTを用いると商品説明と商品レビューとの類似度に差が生まれ、商品レビューを商品説明ごとに分けられないかと考えた。また、冗長性を減らすことで多くの意見を得られるのではと考えた。本研究では、BERTの事前学習モデルを用いて商品説明と商品レビューをベクトル化し、類似度をもとに商品説明に関係する商品レビューを抽出し、MMRを用いて冗長性を減らして表示する手法を提案する。

キーワード レビュー抽出、MMR、BERT、自然言語処理

1 はじめに

通販サイトで商品を購入するとき、多くの人は商品説明から特徴を確認し、商品レビューから実際に使用してどのような感想を持ったのか、何か問題があるかなどを確認する。しかし、多くの通販サイトの商品レビューは新しい順や参考になる商品レビュー順、評価順などで並び替えられ表示されている。これらの並び替え方では、知りたい情報を得るために多くの商品レビューを読み、知りたい情報を探さなければならない。そこで、商品説明ごとに関係する商品レビューを表示すれば、すぐに商品説明について購入者の意見を知ることができると考えた。また、これまで購入決定に有効な商品レビューの抽出に関する研究は多く行われてきたが、商品説明ごとに商品レビューを分ける研究はほとんどない。そこで本研究では、事前にラベル付や新たにモデルを作成をせず、自動で商品レビューを関係する商品説明ごとに分ける手法を提案する。

自動で商品レビューを関係する商品説明ごとに分けるために、文書間類似度を用いた。文書間類似度を用いる理由は、関係する商品説明と商品レビューでは同じ単語や似た表現が使われているので、文書間類似度は高く、関係しない商品説明と商品レビューでは異なる事柄について書かれているので、文書間類似度は低くなり、商品説明ごとに商品レビューを分けられるだけの差が文書間類似度に出るのではないかと考えたからである。

事前実験として商品説明と商品レビューの文書間類似度を用いて、商品レビューを類似度が最も高い商品説明ごとに分けた。結果、商品説明に関係する内容が少なく参考にならない商品レビューや商品説明に関係しない商品レビューがあり、これらの商品レビューが誤って分けられないようにする必要があることがわかった。また、同じ内容の商品レビューが多くあり、少ない商品レビューで多くの情報を得るためには冗長性を減らす必要があることが分かった。誤って商品レビューを分けないようにするために、同じ商品カテゴリーの他の商品の商品レビュー

から、商品説明ごとに関係する内容を多く含む商品レビューと、商品説明に関係しない商品レビューを自動で収集し用いることを考えた。収集した商品レビューをサンプルレビューとし、サンプルレビューと商品説明ごとに分けたい商品レビューとの類似度を用いて、商品レビューを関係する商品説明ごとに分ける。こうすることで、商品説明に関係する内容が多く含まれている商品レビューは商品説明と関係があり、商品説明に関係しないサンプルレビューとの類似度が最も高い商品レビューはどの商品説明とも関係しないと判断されるようにした。また、冗長性を減らすためにMMRを用いて並び替えることを考えた。MMR (Maximal Marginal Relevance) [1] とは、検索クエリの関係性によって並んでいる文書を関連新規性を考慮するように並び替える手法のことである。このMMRを用いて類似度順に並んでいる商品レビューを冗長性を減らすように並び替え、少ない商品レビューで多くの情報を得られるようにする。

本研究では、「リクライニングチェア」の商品レビューを「座り心地」、「カラー」、「リクライニング」に関する3種類の商品説明ごとに分ける実験を行った。行った実験は3種類ある。1つ目は、自動でサンプルレビューを作成するとき、商品説明に関係しないサンプルレビューや複数の商品説明の内容を含む商品レビューが商品説明に関係するサンプルレビューの中にくっ含まれてしまうかを確認する実験である。2つ目は、サンプルレビューを用いて商品レビューを関係する商品説明ごとに分ける実験である。3つ目は、商品説明ごとに分けた商品レビューをMMRを用いて並び替える実験である。

3つの実験の結果、サンプルレビューやMMRを用いたことで、説明写真に関係しない商品レビューの抽出数が減少し、説明写真ごとに購入決定に有効なレビューを冗長性を減らして表示することができた。

本論文における貢献は以下の通りである。

- 同じカテゴリーの他の商品レビューとMMRを用いることによる、説明写真に関係しない商品レビューの抽出数の減少と購入決定に有効な商品レビューの抽出数の増加を確認した

2 関連研究

中野ら [2] は、商品の機能や特徴とそれに対する購入者の意見から構成する、属性-意見のペアを商品レビューの文章から抽出する手法を提案している。品詞や係り受けを用いて素性ベクトルを作成し、SVM を用いて文節に属性や意見が含まれているか否かを分類している。また属性-意見のペアの集合を要約し、要約結果を提示するインターフェースも提案している。

また、佐々木ら [3] は、購入の意思決定を支援するために有効な商品レビューを判別する手法を提案している。この研究では、商品レビューの有効性の判定を行い有効な商品レビューの特徴を分析し、SVM で用いる素性を定義する。そして、SVM で有効な商品レビューか有効でない商品レビューかを分類している。素性は、品詞やキーフレーズ数、文字数などを用いている。

本研究は、SVM や共起、係り受けを用いず事前学習済みの BERT [4] を用いて、文ベクトルを作成し \cos 類似度をもとに商品レビューを分類している。よって、訓練データの作成や、商品や商品ジャンルごとに分類器を作成をせずに分類を行える。

松尾ら [5] は、商品の評価表現を含む文の指標値を算出し、指標値をもとに重要文を抽出する手法を提案している。指標として、評価の極性値、重回帰分析を用いた t 値、文中の評価を表す単語の割合を用いている。指標の値の合計値、または libSVM を用いて重要文を決定する。この研究では商品ジャンルや、評価属性、評価属性を表す語を事前に決める必要がある。

また、平山らは [6]、共起と係り受けを用いた評価属性とその関係性、極性、評価属性に関する文章を抽出している。そして、抽出した評価情報を可視化するシステムを提案している。この研究では、評価属性は商品レビューから抽出している。

本研究は、評価属性を商品レビューから抽出したり事前に決めず、商品説明を評価属性とし、評価属性ごとに商品レビューを表示する。

3 提案手法

商品説明とその商品のカテゴリ名を商品名に含んでいる商品レビューを用いて、商品レビューを関係する商品説明ごとに分ける。以下に提案手法の手順を示す。カテゴリ名を商品名に含んでいる商品レビューの中から、商品説明ごとに関係するサンプルレビューと、どの商品説明とも関係しないサンプルレビューを作成する。サンプルレビューを用いて商品説明に關係する商品レビューを抽出し、MMR を用いて冗長性を減らして表示する。

手順中の BERT は Sentence-BERT [7] の日本語モデル [8] を用いて文ベクトルを作成する。

- (1) 商品説明ごとのサンプルレビューと、どの商品説明とも関係しないサンプルレビューを作成する
 - (a) 商品カテゴリ名を商品名に含んでいる商品レビューを収集する。
 - (b) 収集した商品レビューを文ごとに分割し商品レビュー

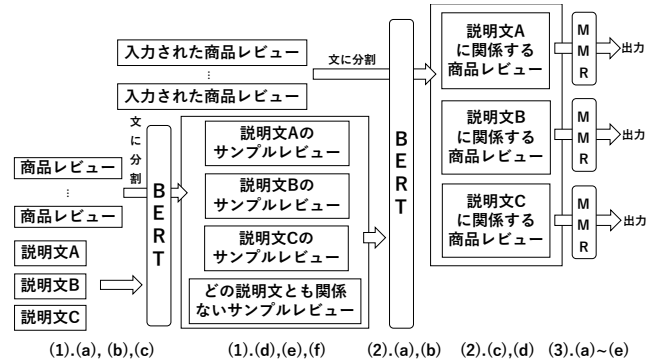


図 1 提案手法の流れ

文とする。

- (c) 商品説明と商品レビュー文を BERT に入力し文ベクトルを作成する。
 - (d) 商品レビュー文と全ての商品説明との \cos 類似度を求め、最も類似度が高い商品説明ごとに分ける。
 - (e) 商品説明ごとに分けられた商品レビュー文を類似度の降順に並べ、閾値を基準に商品説明に關係するサンプルレビューと商品説明に關係しないサンプルレビューに分ける。
 - (f) 商品説明に關係するサンプルレビューの中で、複数の商品説明に關係する、または文中に商品説明に關係しない内容が含まれているレビューを除去する。商品説明ごとに分けられている、商品説明に關係しないサンプルレビューは一つにまとめる。
- (2) 入力された商品レビューを商品説明ごとに分ける
 - (a) 入力された商品レビューを文に分割し入力された商品レビュー文とする。
 - (b) 入力された商品レビュー文とサンプルレビューを BERT に入力し文ベクトルにする。
 - (c) 入力された商品レビュー文と全てのサンプルレビューで \cos 類似度を求める。
 - (d) どの商品説明とも關係しないサンプルレビューとの \cos 類似度よりもいずれかの商品説明のサンプルレビューとの \cos 類似度の方が高いとき、その商品説明に關係する商品レビュー文とする。
 - (3) 商品説明ごとに、商品レビュー文を MMR を用いて冗長性を減らして出力する。
 - (a) 元々の商品レビューが同じかつ、同じ商品説明に分けられているテスト商品レビューを結合し直す。
 - (b) 結合し直した商品レビューを BERT に入力し文ベクトルにする。
 - (c) 結合し直した商品レビューと分けられた商品説明のサンプルレビューとの \cos 類似度を計算する
 - (d) 結合し直した商品レビュー同士の \cos 類似度を計算する
 - (e) (c), (d) の \cos 類似度を用いて MMR で並び替える。

3.1 サンプルレビューの作成

サンプルレビューは商品説明に関係するものと関係しないものの2種類を作成する。

商品説明に関係するサンプルレビューは、商品レビューには一つの観点でも様々な意見があるが、意見ごとに類似度に差ができてしまうことがあり、類似度に差が生まれないようにするために用いる。例えば、座り心地に関する商品説明と「座り心地がとてもよく頭までつけることができリラックスできます」との \cos 類似度は 0.735 だが、「座った感じはもうちょっと柔らかいのかな〜とと思っていましたが結構硬め」との \cos 類似度は 0.461 と差ができています。このため MMR を用いるときに後者の商品レビューが下の方に並び替えられてしまうことがある。そこで、一つの観点でも様々な意見の文を複数収集し、その文との類似度を用いることで類似度に差が起らないようにする。

商品説明に関係しないサンプルレビューは、商品説明ごとに商品レビューを分けたとき、商品説明に関係しない商品レビューを含んでいると MMR を用いて並び替えるときに上位に並び替えられてしまうことがあるので、どの商品説明にも関係しない商品レビューを関係しないと判断させるためのものである。

商品説明に関係するサンプルレビューを作成するとき、商品説明に関係しないサンプルレビューが含まれてはいけない。しかし、複数の観点について書かれている商品レビューが多く存在する。複数の観点が含まれている状態で抽出を行っても別の観点の内容が含まれてしまうので、商品レビューを「。」と空白で分割する。しかし、分割しても複数の観点を含んでいる文が存在し、このようなサンプルレビューは除去する必要がある。

複数の観点を含んでいるサンプルレビューを除去するために、複数の観点を含んでいる文は複数の商品説明で分けられていると考えた。また、複数の商品説明の内容を含んでいる文章は1つの商品説明の内容を含んでいる文より、同じ複数の商品説明の内容を含んでいる文との類似度の方が高くなると考えた。そこで全ての商品説明のペアで以下の手順で複数の商品説明の内容を含んでいるサンプルレビューを削除した。全ての商品説明のペアは商品説明の数を n としたとき ${}_n C_2$ 組みあるので、全ての組みで以下の手順を行う。

- (1) 商品説明のペアをそれぞれ A, B とし、A に分けられているサンプルレビューと A の商品説明の \cos 類似度を計算する。
- (2) A に分けられているサンプルレビューと B に分けられている全てのサンプルレビューとの \cos 類似度を計算する。
- (3) 同様に B に分けられているサンプルレビューも B の商品説明と A に分けられている全てのサンプルレビューとの \cos 類似度を計算する。
- (4) A に分けられているサンプルレビューの中で A の商品説明との \cos 類似度よりも B に分けられているサンプルレビューとの \cos 類似度の方が高いとき、そのサンプルレビューを削除する。

- (5) 同様に B に分けられているサンプルレビューも削除する。

このようにして、複数の観点を含んでいるサンプルレビューを除去し、残ったサンプルレビューを商品説明に関係するサンプルレビューとする。

3.2 入力された商品レビューを商品説明ごとに分割

3.1 節で作成した2種類のサンプルレビューを用いて、入力された商品レビューを商品説明ごとに分ける。手順としては BERT を用いてサンプルレビューと入力された商品レビューから文ベクトルを作成し、 \cos 類似度を計算、 \cos 類似度が高いサンプルレビューが属している商品説明に分けるという流れである。3.1 節と異なる点は、商品レビューが必ずどれかの商品説明に分けられるのではなく、どの商品説明とも関係しないサンプルレビューとの \cos 類似度が最も高いときは商品説明に分けられないという点である。これによって、商品説明に関係する商品レビューのみが商品説明に分けられるようにする。

3.3 MMR を用いた並び替え

3.2 節で商品説明ごとに分けられた商品レビューを MMR を用いて冗長性を減らして表示する。本研究で用いる MMR の式は以下である。

まだ並び替えられていない商品説明ごとに分けられた商品レビューの文ベクトル D_i に対して

$$MMR_score = \operatorname{argmax}[\lambda \times Sim_1(D_i, S_{max}) - (1 - \lambda) \times Sim_2(D_i, D_j)]$$

$Sim_1 = \cos(D_i, S_{max})$, $Sim_2 = \cos(D_i, D_j)$ とする。

D_j は既に並び替えられた商品説明ごとに分けられた商品レビューの文ベクトル、 S_{max} は D_i の商品レビューと最も類似度が高いサンプルレビューの文ベクトルである。 Sim_1 は D_i と S_{max} の \cos 類似度であり、 Sim_2 は D_i と D_j の類似度である。 λ は Sim_1 と Sim_2 のどちらを重要視するか、つまり商品説明に関係することを優先するか、既に並び替えられている商品レビューとの \cos 類似度の大きさを優先するかを決める値である。本研究では両方を同等に重要と考え $\lambda = 0.5$ としている。

最初に並び替えられる商品レビューは、既に並び替えられている商品レビューがまだ無いので MMR_score が計算できない。そこで本研究では、サンプルレビューとの \cos 類似度が最も高い商品レビューを最初に並べ、それ以降は MMR_score が最も高い商品レビューから順に並べていく。

手順3の流れは、まず商品レビュー文を結合し直す。理由としては、商品レビューの中には「ブラウンを購入しました。やや濃い色見ですが落ち着いていて気に入っています。」のように「。」や空白で分割すると何に対しての商品レビューかわからなくなってしまうことがあるからである。次に、結合し直した商品レビューを BERT に入力し文ベクトルを作成する。この文ベクトルを用いて MMR に必要な結合し直した商品レビューと商品説明ごとのサンプルレビューとの \cos 類似度と、結合し直した商品レビュー同士の \cos 類似度を計算する。最後にこの2種類の \cos 類似度を用いて MMR を計算し並び替えて表示する。

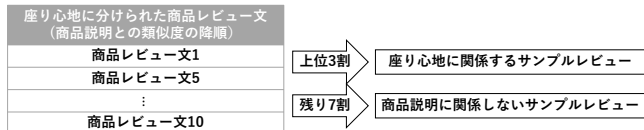


図 2 閾値が 0.3 のときのサンプルレビューの分割例

4 実 験

本実験では、「リクライニングチェア」の商品レビューを「座り心地」、「カラー」、「リクライニング」の商品説明ごとに分け、冗長性を減らすように並び替えることを目的とする。4.2 節の実験では、商品説明に関するサンプルレビューの中に商品説明に関係しないサンプルレビューや複数の商品説明の内容を含む商品レビューがいくつ含まれるか確認する。4.3 節の実験では、サンプルレビューを用いて 6,946 件の商品レビューを商品説明ごとに分け、商品説明に関係しない商品レビューがいくつ含まれるか確認する。4.4 節の実験では、サンプルレビューを用いて 1,438 件の商品レビューを商品説明ごとに分け、商品説明に関係しない商品レビューがいくつ含まれるか確認する。4.5 節の実験では、4.3 節、4.4 節の実験で商品説明ごとに分けた商品レビューを MMR を用いて並び替え、上位 100 件の商品レビュー中の商品説明に関係しない商品レビューと購入決定に有効な商品レビューの数を確認する。

全ての実験は、サンプルレビューの作成時に用いる閾値を変更しながら行う。閾値は 0.1~0.9 の 0.1 刻みで変更し、3 章の (1).(d) で商品説明ごとに分けた商品レビュー文を、商品説明に関するサンプルレビュー：商品説明に関係しないサンプルレビュー=閾値：1-閾値に分ける。例えば、閾値が 0.3 のときの座り心地の場合、図 2 のように座り心地に分けられた商品レビュー文を商品説明との cos 類似度の降順に並べ、上位 3 割を座り心地に関するサンプルレビューに、残り 7 割を商品説明に関係しないサンプルレビューとする。

4.1 実験データ

実験には、国立情報学研究所から提供されている楽天データセット¹の商品レビューを利用する。実験に用いるデータは 4 種類あり、サンプルレビューを作成するための商品レビュー、入力に用いる商品説明に書かれている商品説明、実際に商品説明ごとに分ける 2 種類の商品レビューである。

このデータセットから 2,015 年 2 月から 2,015 年 4 月の間の商品名に「リクライニングチェア」が含まれている 3,569 件の商品レビューを「。」と空白で分割したときの 13,465 件をサンプルレビューを作成するための商品レビューとする。

入力に用いる商品説明は「リクライニングチェア」の商品に実際に用いられていた、「ボリュームたっぷりのヘッドレストは、厚み約 20cm あり、座った瞬間から優しく包まれるような安心感ある心地よさ。腰部分は疲れにくいランバーサポート形

状設計で、頭から足先まで贅沢な座り心地です。」、「大人ムード感じるブラウンシックで落ち着きがあり、高級感のある人気のカラー。」、「座面に腰掛けて背もたれに体重をかけるだけでらくらくリクライニング。一体型のオットマンは、リクライニングすると連動してせり上がります」の 3 種類を用い、以降「座り心地」、「カラー」、「リクライニング」とする。

入力する商品レビューの数が精度に影響するか確かめるために、商品名に「リクライニングチェア」が含まれている 452 件と 1,979 件の商品レビューを「。」と空白で分割した 1,438 件と 6,946 件をテスト商品レビューとする。そして、「座り心地」、「カラー」、「リクライニング」それぞれに対し、関係するを正解、関係していないを不正解として一つずつ読みラベル付けを行う。6,946 件のテスト商品レビューには「座り心地」：1,377 件、「カラー」：682 件、「リクライニング」：478 件、1438 件のテスト商品レビューには「座り心地」：340 件、「カラー」：137 件、「リクライニング」：44 件の正解ラベルがついている。

4.2 サンプルレビューを作成する実験

この実験では、3 種類の商品説明に対して閾値を基準にサンプルレビュー作成用の商品レビューを分ける。そして、商品説明に関するサンプルレビューの中で、商品説明に関係しない内容も含まれている商品レビュー、商品説明に関係しない内容のみの商品レビューを除去したときの精度の確認をする。

4.2.1 結果・考察

表 1 に実験結果を示す。商品説明に関するサンプルレビューは閾値が 0.4~0.8 のときは、ほとんど同じ商品レビュー文であった。商品説明に関係しない内容も含んでいるサンプルレビューは、「座り心地」では、値段に対しての内容が含まれているサンプルレビュー、「カラー」では、商品の全体的な評価が書かれているサンプルレビュー、「リクライニング」では、サイズに対しての内容が含まれているサンプルレビューであった。商品説明に関係しない内容のみのサンプルレビューは、「座り心地」の入力した商品説明には「ヘッドレスト」について書いてあるが「フットレスト」に関するサンプルレビューがいくつ含まれていた。「リクライニング」はオットマンの組み立てや座高の変更についてのサンプルレビューが含まれていた。

実験の結果より、商品説明に関するサンプルレビューは抽

表 1 サンプルレビューの作成結果

閾値 の値	座り心地			カラー			リクライニング		
	関係 ある	複数 内容	関係 ない	関係 ある	複数 内容	関係 ない	関係 ある	複数 内容	関係 ない
0.1	20	1	2	88	8	0	25	2	0
0.2	21	1	3	80	1	0	29	1	3
0.3	22	1	4	69	1	0	30	1	5
0.4	23	1	5	60	1	0	33	1	8
0.5	22	1	4	59	1	0	35	1	9
0.6	23	1	5	57	1	0	35	1	8
0.7	23	1	5	55	1	0	35	1	8
0.8	23	1	5	53	1	0	35	1	8
0.9	23	1	5	53	1	0	39	1	12

1：楽天グループ株式会社 (2014)：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)。https://doi.org/10.32130/idr.2.0

出できるが、誤ったサンプルレビューも含まれてしまうことがわかった。しかし閾値によっては、商品説明とは関係しないサンプルレビューの中に値段やサイズなどの商品説明に関係しない内容のみのサンプルレビューが含まれているので、商品レビューを商品説明ごとに分けるときに商品説明に関係しない内容のみの商品レビューが抽出されるとは考えにくい。

4.3 6,946 件の商品レビューを商品説明に分ける実験

この実験では、閾値ごとに作成したサンプルレビューを用いて 6,946 件のテスト商品レビューを「座り心地」、「カラー」、「リクライニング」の 3 種類の商品説明ごとに分ける。この実験では 2 つの指標を見ていく。

1 つ目は商品説明ごとに閾値を変化させたときの以下の式で表す適合率である。

$$Precision = \frac{T}{T+F}$$

T 類似度の降順に抽出した商品説明に関する商品レビュー数
F T に含まれる最小の類似度の商品レビューよりも類似度が高い、商品説明に関係しない商品レビュー数

類似度を基準にする理由は、MMR で類似度を用いるとき、商品説明に関係しない商品レビューが類似度が高いと、上位に並び替えられてしまうからである。そのため、商品説明に関する商品レビューの抽出数を増加させ、商品説明に関係しない商品レビューが類似度順のどこに含まれているのか確認する。

2 つ目は、商品説明に関する商品レビューの抽出率である。適合率が高くても抽出数が少なすぎると商品レビューの内容が偏ってしまう可能性があるからである。

4.3.1 結果・考察

表 2, 表 4, 表 6 の 1 行目は、サンプルレビューを用いずテスト商品レビューを商品説明との cos 類似度の最も高い商品説明に分けたときの結果である。同様にサンプルレビューを用いない場合の抽出率は表 3, 表 5, 表 7 の 1 列目に示す。適合率の結果は縦軸が閾値の値、横軸が抽出する商品説明に関する商品レビュー数である。また、() 中の値は分けられた全ての商品説明に関する商品レビューを抽出したときの適合率である。

「座り心地」に関して、表 2, 表 3 に示す。サンプルレビュー

表 2 座り心地の適合率

	200	400	600	800	1000	1200	1400
-	0.975	0.957	0.910	0.848	0.722	0.509	(0.354)
0.1	(0.972)	-	-	-	-	-	-
0.2	(0.975)	-	-	-	-	-	-
0.3	0.975	(0.958)	-	-	-	-	-
0.4	0.966	0.928	(0.920)	-	-	-	-
0.5	0.961	0.934	(0.891)	-	-	-	-
0.6	0.961	0.928	0.884	(0.805)	-	-	-
0.7	0.961	0.928	0.884	0.799	(0.759)	-	-
0.8	0.961	0.925	0.890	0.808	0.702	(0.688)	-
0.9	0.961	0.925	0.892	0.804	0.708	(0.589)	-

表 3 座り心地の抽出率

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
-	0.906	0.024	0.084	0.198	0.312	0.429	0.555	0.641	0.747	0.803

を用いないときの商品説明に関する商品レビューの抽出率は 0.906 と十分に抽出できている。しかし 1,000 件の商品説明に関する商品レビューを抽出するときの適合率は 0.722 と高かったが、それ以降急激に下がり最終的に 0.354 となった。また、cos 類似度が低い商品説明に関する商品レビューは商品説明に関わる内容が一部分であり大部分は関係しない内容であった。サンプルレビューを用いないときはそのような商品レビューも関係する商品説明ごとに分けることができるが、商品説明に関係ない商品レビューも多く含まれる結果となった。

サンプルレビューを用いる場合、閾値の値を上げると商品説明に関する商品レビューの抽出率は増えるが最終的な適合率は下がる傾向になった。これは、閾値の値が小さくなるほど商品説明との cos 類似度が高かった商品レビュー文が商品説明に関するかに関わらず、商品説明に関係しないサンプルレビューに含まれるようになるので、一部分に商品説明に関わらない内容が含まれている商品レビューや商品説明との cos 類似度が高いが商品説明に関係しない商品レビューが、商品説明とは関係しないと判断されるようになるからだと考える。

「座り心地」に関して、閾値の値が 0.5 のときは商品説明に関する商品レビューの抽出率が 0.429 で適合率が 0.891, 0.6 のときの商品説明に関する商品レビューの抽出率が 0.555 で適合率が 0.805 となった。抽出できなかった商品説明に関する商品レビューは商品説明に関わる内容が多い商品レビューが多く、抽出できた商品説明に関する商品レビューは商品説明に関わる内容のみであったり一部分に商品説明に関係しない内容が含まれている商品レビューであった。他の閾値では、商品説明に関する商品レビューの抽出率が低かったり、商品説明に関係しない商品レビューを多く含んでいるので、商品説明ごとに商品レビューを分けられているとは言えない結果だった。

「カラー」に関して、表 4, 表 5 に示す。サンプルレビューを用いないときの商品説明に関する商品レビューの抽出率は 0.947 であった。しかし、適合率は 0.316 と小さい値となり、300 件の抽出でも適合率は 0.561 と小さい値であった。

サンプルレビューを用いると、閾値によっては 300 件の抽出

表 4 カラーの適合率

	100	200	300	400	500	600	700
-	0.869	0.740	0.561	0.442	0.385	0.341	(0.316)
0.1	0.980	(0.977)	-	-	-	-	-
0.2	0.980	(0.978)	-	-	-	-	-
0.3	0.980	(0.973)	-	-	-	-	-
0.4	0.980	0.970	(0.963)	-	-	-	-
0.5	0.970	0.952	(0.947)	-	-	-	-
0.6	0.980	0.947	0.928	(0.921)	-	-	-
0.7	0.980	0.952	0.914	(0.890)	-	-	-
0.8	0.980	0.952	0.884	0.833	(0.832)	-	-
0.9	0.980	0.904	0.854	0.756	(0.689)	-	-

表 5 カラーの抽出率

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
-	0.947	0.193	0.265	0.272	0.307	0.373	0.467	0.521	0.604	0.674

でも適合率が0.90以上のときもあり、サンプルレビューの有無で精度に差が出た。サンプルレビューを用いないときに商品説明とのcos類似度の降順に並べて類似度を見てみると、商品説明に関係する商品レビューのcos類似度が低い傾向があった。これは、商品説明と商品レビュー文と異なる文とのcos類似度を求めているのでうまくいかないことがあることが原因だと考える。逆にサンプルレビューを用いたときは商品説明に関係する商品レビューのcos類似度が高い傾向があった。このことより、商品説明に関係するサンプルレビューは商品説明に関係する商品レビューを分けるときに有効であることがわかった。

また「カラー」は、閾値が0.4~0.9のときは商品説明に関係する商品レビューの抽出率は若干少なくなっているが、最終的な適合率は0.9の時を除き0.800以上となった。適合率が高いのは、「カラー」は他の商品説明とは異なり、商品説明と関係しないサンプルレビューが1つも無いことが原因だと考える。

「カラー」は閾値が0.6のときの商品説明に関係する商品レビューの抽出率が0.467で適合率が0.921、0.7のときの抽出率が0.521で適合率が0.890となった。他の閾値は、商品説明に関係する商品レビューの抽出率が低かったり、商品説明に関係しない商品レビューを多く含んでしまっている。

「リクライニング」に関して、表6、表7に示す。サンプルレビューを用いないとき、抽出率は0.430、適合率は0.20となった。これは、商品レビューの特徴として「リクライニング」に関する内容は「座り心地」の内容とともに書かれていることがあることが原因だと考える。サンプルレビューは「リクライニング」に関することしか書かれていないものなので、抽出する商品説明に関係する商品レビューは「座り心地」に関する内容が含まれていないものが多くなっている。「リクライニング」と「座り心地」の両方が書かれている商品説明に関係するの多くは「座り心地」に分けられていた。これはサンプルレビューを用いたときでも少なからず影響を受けている結果となった。

また他の「座り心地」や「カラー」と比較すると適合率が低くなった。これは、「リクライニングするときのレバーが重い」、「椅子が重たく移動しづらい」などの同じ言葉だが対象が異なる商品レビューを抽出してしまうことが原因であると考えられる。

表6 リクライニングの適合率

	30	60	90	120	150	180	210
-	0.937	0.895	0.731	0.662	0.568	0.375	(0.200)
0.1	(0.916)	-	-	-	-	-	-
0.2	(0.956)	-	-	-	-	-	-
0.3	1.00	(0.975)	-	-	-	-	-
0.4	1.00	(0.980)	-	-	-	-	-
0.5	1.00	0.953	(0.881)	-	-	-	-
0.6	1.00	0.923	0.841	(0.789)	-	-	-
0.7	1.00	0.923	0.858	(0.696)	-	-	-
0.8	1.00	0.923	0.825	0.677	(0.531)	-	-
0.9	1.00	0.909	0.810	0.655	0.485	(0.396)	-

表7 リクライニングの抽出率

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.430	0.023	0.046	0.081	0.104	0.140	0.211	0.244	0.299	0.343

サンプルレビューを用いた場合でも、商品説明に関係する商品レビューの抽出率は他の商品説明に比べてかなり低くなった。原因はサンプルレビューがないときと同様だと考える。

「リクライニング」に関しては、閾値でみると、0.6のときが商品説明に関係する商品レビューの抽出率も他の閾値よりも相対的に良く、適合率も0.789となったが、「座り心地」や「カラー」と比べてよくない結果となった。

4.4 1,438 件の商品レビューを商品説明に分ける実験

この実験は同様に1,438件のテスト商品レビューを「座り心地」、「カラー」、「リクライニング」の3種類の商品説明に分ける。この実験でも同様に商品説明に関係する商品レビューの抽出率と適合率の2つの指標を見ていく。

4.4.1 結果・考察

4.3節の実験と同様にサンプルレビューを用いないときの結果は、表8、表10、表12の1行目に示す。縦軸が閾値の値、横軸がとってくる商品説明に関係する商品レビューの数である。また、サンプルレビューを用いない場合の抽出率は表9、表11、表13の1列目に示す。括弧の中に書かれている値は、その閾値で分けたときの分けられた全ての商品説明に関係する商品レビューを抽出したときの適合率である。

表8 座り心地の適合率

	50	100	150	200	250	300
-	1.00	1.00	0.955	0.781	0.668	(0.460)
0.1	-	-	-	-	-	-
0.2	(1.00)	-	-	-	-	-
0.3	(1.00)	-	-	-	-	-
0.4	0.980	(0.955)	-	-	-	-
0.5	0.980	(0.959)	-	-	-	-
0.6	1.00	0.952	(0.924)	-	-	-
0.7	1.00	0.970	0.937	(0.892)	-	-
0.8	1.00	0.980	0.925	0.833	(0.820)	-
0.9	1.00	0.980	0.925	0.840	(0.691)	-

表9 座り心地の抽出率

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.832	0	0.035	0.111	0.191	0.279	0.391	0.511	0.617	0.705

表10 カラーの適合率

	20	40	60	80	100	120	140
-	0.952	0.975	0.779	0.465	0.348	0.263	(0.229)
0.1	1.00	(1.00)	-	-	-	-	-
0.2	1.00	1.00	(1.00)	-	-	-	-
0.3	1.00	1.00	(1.00)	-	-	-	-
0.4	1.00	1.00	(0.981)	-	-	-	-
0.5	1.00	1.00	(0.967)	-	-	-	-
0.6	1.00	1.00	0.967	(0.935)	-	-	-
0.7	1.00	1.00	0.967	0.898	(0.900)	-	-
0.8	1.00	1.00	0.952	0.888	(0.815)	-	-
0.9	1.00	1.00	0.952	0.851	0.675	(0.660)	-

表11 カラーの抽出率

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.912	0.270	0.313	0.321	0.379	0.437	0.518	0.591	0.678	0.766

「座り心地」に関して、表 8, 表 9 に示す。サンプルレビューを用いないときの商品説明に関する商品レビューの抽出率は 0.832, 適合率は 0.460 と 4.3 節の実験結果と差はなかった。

しかしサンプルレビューを用いたときは、適合率の傾向は変わらなかったが商品説明に関する商品レビューの抽出率が全体的に下がった。これは「座り心地はいいです」などあまり参考にならないテスト商品レビューが 4.3 節の実験の場合より割合が高く、商品説明に関係しないと判断されたからだと考える。

「カラー」に関して、表 10, 表 11 に示す。サンプルレビューを用いないときは、商品説明に関する商品レビューの抽出率は 0.912 だが、適合率は 80 件の商品説明に関する商品レビューの抽出以降、急に下がり 0.229 となった。閾値を変化させた場合は、評価指標は多少の誤差がある程度で、商品説明に関する商品レビューの抽出率は若干上がった。「カラー」に関して商品レビュー数によって精度が悪くなることはなかった。

「リクライニング」に関して、表 12, 表 13 に示す。サンプルレビューを用いないとき商品説明に関する商品レビューの抽出率は 0.318, 適合率は 0.278 とどちらも低い値となった。

閾値を変化させても、0.9 以外は適合率は高いが商品説明に関する商品レビューの抽出率は低かった。

「リクライニング」の商品説明に関する商品レビューの数は 44 件ともともと少なかった。少ない場合でも全く抽出できないことはないが、改善する必要があることがわかった。

実験結果から、商品レビューの数ではなく、商品説明に関する商品レビューの数が影響することがわかった。

4.5 MMR を用いて並び替える実験

この実験では、4.3 節, 4.4 節の実験で商品説明ごとに分けた商品レビューを MMR を用いて冗長性を減らすように並び替える。そして、並び替えた上位 100 件中の「購入決定に有効な商品レビューの数」、「商品説明に関係しない商品レビューの数」を検証する。購入決定に有効な商品レビューとは、「ただリクライニングの動きが少し固いかな」のような商品レビューのことであり、表 14, 表 15 の「A」に、商品説明に関係しない商品レビューの数は表 14, 表 15 の「B」に示す。また、閾値によ

ては 100 件抽出できない場合があり、それぞれの抽出数も示す。

4.5.1 結果・考察

以下に実験結果を示す。6,946 件のテスト商品レビューのとき、サンプルレビューを用いない場合、「座り心地」は他の「カラー」や「リクライニング」に比べて商品説明に関係しないテスト商品レビューは少なかった。しかし、結合したことによって商品説明に関するテスト商品レビューに関係しない内容が結合されることが多くあり、購入決定に有効な商品レビューを表示できているとは言えない結果であった。

「カラー」の関関係しないテスト商品レビューは、サイズに関するものや、どの商品説明にも関係しないものが多かった。また、「色はボールドです」のように色についてのみの内容のものもいくつかあったが、テスト商品レビューを結合することで、何色を購入したかと一緒に実際にその色はどうだったか書かれているものも多くあった。「カラー」に関して、ある色に対して詳しく書かれている商品レビューを購入決定に有効な商品レビューとしたが、多くの商品レビューは「色は落ち着いていて、お洒落です」のように、どの色に対しての書かれているかわからないことがあった。これらの商品レビューは購入決定に有効な商品レビューとはしなかったので、購入決定に有効な商品レビューの数が少ない結果となった。だが、これらの商品レビューも購入決定に有効な商品レビューに含めると、他の商品説明と購入決定に有効な商品レビューの数は変わらなかった。

「リクライニング」は関係しないテスト商品レビューが 57 件と過半数を上回っている。これは、類似度順に並べたときに上位に商品説明に関係しない商品レビューが多いことが原因だと考える。また、購入決定に有効な商品レビューの数も少なく、並び順も商品説明に関係しない商品レビューが上位に並べられているのでうまくいかなかった。

サンプルレビューを用いた場合は、どの商品説明でも閾値が小さくなるにつれて商品説明に関係しない商品レビューの抽出数が減少した。しかし閾値が小さいと、抽出数が少なかったり、同じ内容の購入決定に有効な商品レビューを多く抽出した。逆に閾値が大きいと、商品説明に関係しない商品レビューや、商品説明に関係しない内容を多く含む商品レビューが多く抽出された。実験の結果、閾値が 0.5 や 0.6 のときに 3 種類の商品説明に共通して、商品説明に関係しない商品レビューの抽出数が少

表 12 リクライニングの適合率

	2	4	6	8	10	12	14
-	1.00	1.00	1.00	1.00	0.833	0.600	(0.378)
0.1	-	-	-	-	-	-	-
0.2	-	-	-	-	-	-	-
0.3	-	-	-	-	-	-	-
0.4	(1.00)	-	-	-	-	-	-
0.5	(0.500)	-	-	-	-	-	-
0.6	0.666	0.800	0.857	0.888	(0.900)	-	-
0.7	0.666	0.800	0.750	0.800	0.833	(0.785)	-
0.8	0.666	0.800	0.750	0.800	0.833	(0.785)	-
0.9	0.666	0.800	0.750	0.727	0.714	(0.687)	-

表 13 リクライニングの抽出率

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
-	0.318	0	0	0.022	0.022	0.204	0.250	0.250	0.250

表 14 6,946 件のテスト商品レビューに MMR を用いた結果

	座り心地			カラー			リクライニング		
	抽出数	A	B	抽出数	A	B	抽出数	A	B
-	100	47	15	100	41	30	100	38	57
0.1	37	37	0	100	43	0	12	13	0
0.2	100	82	0	100	52	0	23	23	1
0.3	100	84	2	100	52	0	39	38	1
0.4	100	84	1	100	56	0	50	49	1
0.5	100	84	1	100	57	1	73	67	4
0.6	100	83	5	100	56	1	100	84	14
0.7	100	75	5	100	52	9	100	77	21
0.8	100	66	8	100	50	6	100	62	29
0.9	100	64	12	100	43	10	100	53	37

なく、異なる内容の購入決定に有効な商品レビューを抽出した。

1,979 件のテスト商品レビューのとき、サンプルレビューを用いない場合は 6,946 件のテスト商品レビューのとき同様、関係しない商品レビューが多く、うまくいかなかった。

サンプルレビューを用いると、「座り心地」は閾値が 0.2~0.4 のとき、商品説明に関係しない商品レビューはほとんどないが、「座り心地はいいです」のような商品レビューが多く表示され、内容がわからないわけではないが参考にならない商品レビューが多かった。閾値が 0.5~0.7 のときは、そのような商品レビューは表示されなかったが、閾値が 0.8,0.9 のときは商品説明に関係しないテスト商品レビューが多く表示された。

「カラー」も閾値が小さくなるにつれ、抽出数が下がっていった。抽出されなくなった商品レビューは一部分に関係しない内容を含んでいることが多く、閾値が小さくても「色もきれいで」のような商品レビューが多く抽出された。

「リクライニング」に関しては、どの閾値でも商品レビューの抽出数が少ないので MMR を用いても、商品説明に関係する商品レビューと関係しない商品レビューが混合して並んでいた。

実験結果より、6,946 件と 1,979 件の両方で、サンプルレビューを用いた方が商品説明に関係する商品レビューを多く抽出できた。しかし、閾値が大きすぎたり小さすぎたり、そもそもの商品説明に関係する商品レビューの数が少ないとうまくいかないことがわかった。

5 おわりに

本研究は、商品レビューを関係する商品説明ごとに分け冗長性を減らして表示することを目的としている。

本研究では、「リクライニングチェア」の商品の商品レビューを「座り心地」、「カラー」、「リクライニング」の 3 種類の商品説明に分ける実験を行った。実験として、提案手法のサンプルレビューを作成したときの精度や、商品説明ごとに商品レビューを分けたときの精度を確認しつつ、MMR を用いて並び替えたときの上位 100 件の商品レビュー中の商品説明に関係しない商品レビューと購入決定に有効な商品レビューの数を確認した。

実験の結果、452 件と 1,979 件の商品レビュー共通して閾値が 0.1~0.4 のとき、内容が単純で購入決定に有効になりずら

かったり、抽出数が少なく、閾値が 0.7~0.9 のときは商品説明と関係しない商品レビューが多かったり、商品説明と関係ある商品レビューでも関係しない内容が多く含まれていた。しかし、閾値が 0.5, 0.6 のときは、商品説明と関係しない商品レビューの抽出数を減らし、購入決定に有効な商品レビューを多く抽出することができた。また、商品説明に関係する商品レビューの数が少ない場合はうまくいかないことがわかった。実験を通じて、同じカテゴリーの他の商品レビューと MMR を用いることによる、説明写真に関係しない商品レビューの抽出数の減少と購入決定に有効な商品レビューの抽出数の増加を確認した。

今後の課題として、別の商品ジャンルでも同様の 0.5 や 0.6 の閾値と提案手法を用いて商品レビューを関係する商品説明ごとに分けられるかの検証する必要がある。商品ごとに用いる閾値が変わってしまえばいけないので、入力する商品説明、商品レビューを変更し、同様の条件で検証する。また、商品説明に関係する商品レビューが元々少ないときの抽出精度の向上する方法の検討も必要である。商品レビューが多くて知りたい情報が得られないときに、提案手法は役立つが、知りたい情報が書かれている商品レビューが少ないときでも、知りたい情報だけを得られるように改善する必要があると考える。

謝辞 本研究の一部は JSPS 科研費 19H04218 の助成を受けたものです。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グループ株式会社から提供を受けた「楽天データセット」(<https://rit.rakuten.com/data-release/>) を利用しました。

文 献

- [1] Jaime Carbonell and Jade Goldstein. The use of mmr, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 335–336, 1998.
- [2] 中野裕介, 湯本高行, 新居学, 佐藤邦弘ほか. 商品レビュー要約のための属性-意見ペア抽出. 研究報告組込みシステム (EMB), Vol. 2014, No. 15, pp. 1–7, 2014.
- [3] 佐々木優衣, 関洋平. 商品レビューを対象とした有用性の定義と判別. In *DEIM Forum B5-1*, 2014.
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [5] 松尾哉太, 新妻弘崇, 太田学ほか. レビュー文書における重要文選択の一手法. 研究報告データベースシステム (DBS), Vol. 2015, No. 12, pp. 1–7, 2015.
- [6] 平山拓央, 湯本高行, 新居学, 佐藤邦弘ほか. 語の共起と極性に基づく商品レビュー閲覧支援システム. 研究報告データベースシステム (DBS), Vol. 2012, No. 3, pp. 1–9, 2012.
- [7] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*, 2019.
- [8] sonoisa(日鉄ソリューションズ株式会社). 【日本語モデル付き】2020年に自然言語処理をする人にお勧めしたい文ベクトルモデル, 2021.

表 15 1,979 件のテスト商品レビューに MMR を用いた結果

	座り心地			カラー			リクライニング		
	抽出数	A	B	抽出数	A	B	抽出数	A	B
-	100	39	42	100	11	71	100	15	84
0.1	0	-	-	36	13	0	0	-	-
0.2	12	9	0	42	17	0	0	-	-
0.3	38	16	0	44	17	0	0	-	-
0.4	68	33	1	51	20	1	1	1	0
0.5	97	58	1	56	21	1	4	2	2
0.6	100	73	2	68	25	1	12	9	2
0.7	100	61	4	80	28	3	16	10	4
0.8	100	62	13	97	33	8	19	11	7
0.9	100	60	17	100	33	15	34	13	18