

Title	複数の製品を対比する文の自動検出
Author(s)	花岡, 裕
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/12039">http://hdl.handle.net/10119/12039</a>
Rights	
Description	Supervisor: 白井 清昭, 情報科学研究科, 修士

修 士 論 文

**複数の製品を対比する文の自動検出**

北陸先端科学技術大学院大学  
情報科学研究科情報科学専攻

花岡 裕

2014年3月

修 士 論 文

## 複数の製品を対比する文の自動検出

指導教員 白井清昭 准教授

審査委員主査 白井清昭 准教授  
審査委員 島津明 教授  
審査委員 東条敏 教授

北陸先端科学技術大学院大学  
情報科学研究科情報科学専攻

1210044 花岡 裕

提出年月: 2014 年 2 月

## 概要

近年、ユーザレビューを対象とした評判分析は製品の購入を検討するユーザに有益な情報を提供する手段として注目されている。本研究ではユーザによるレビュー文のうち、特に複数の製品を対比した文に着目する。例えば「ipadはnexus7より重い」といった文のようにnexus7とipadを対比している文は、タブレット端末の購入を検討しているユーザにとって有益な情報となる。本論文は、複数の製品を対比した文(以下、製品対比文と呼ぶ)を検出することを目的とする。

本論文の提案手法の流れを以下に示す。まず、ユーザが評判を知りたい対象製品( $P_t$ とする)についてレビュー文を取得する。次に、対象製品 $P_t$ の競合製品リストを作成する。次に、製品対比文の検出を行う。最後に、製品対比文であると判定した文を出力する。

まず、対象製品のレビュー文を取得する。“ $P_t$  レビュー”をクエリとして検索エンジンでウェブ検索を行う。得られたウェブページをHTMLタグ及び句点で分割したものをレビュー文の集合とする。

次に $P_t$ と競合する製品のリストを作成する。競合製品とは、 $P_t$ と同じ種類の製品で、購入を検討する際に比較の対象となりうる製品とする。ここではウェブサイト「価格.com」における $P_t$ を含む製品カテゴリの製品リストをそのまま競合製品リストとする。

次に製品対比文の判定を行う。ここでは、与えられたレビュー文が製品対比文であるか否かを判定する2つの手法を提案した。1つはルールベースの手法、もう1つは機械学習に基づく手法である。

ルールベースの手法では二つのルールを設計した。一つは製品名が複数出現する文を製品対比文と判定するルールAである。ここで製品名は対象製品 $P_t$ もしくは競合製品リスト中の製品とする。これは、製品名が2つ以上現れる文においては、それらの製品が対比されているという考えに基づく。ただし、“iPadもNexus7も十分高性能”のように、製品名が複数現れた場合でも、出現した製品を同列に扱って評価を行っているときなど、それらの製品を対比しているとはみなされない場合がある。そこで、レビュー文を節に分割し、2つ以上の節に異なる製品名が出現しなければならないとした。

もう一つのルールは、人手作成した対比表現辞書を基に対比文の判定を行うルールBである。対比表現辞書は、対比文に特徴的に現れる表現を対比表現とし、これを収集した辞書である。ルールBは、文中に辞書に登録された対比表現が出現したとき、それを製品対比文と判定する。ただし、文中に対比表現が出現するだけでは必ずしも製品対比文であるとは限らない。そこで、対比表現について、それが製品対比文に出現するための付加条件を設定した。

機械学習に基づいた手法では、対比文がタグ付けされた製品レビュー文の集合を訓練データとし、与えられた文が製品対比文か否かを判定する二値分類器を教師あり学習する。本論文で採用した学習アルゴリズムはSVMである。製品対比文の判定に有用であると考えられる以下の9つの学習素性を設計した。(1)文中の自立語。(2)対比表現辞書内の対比

表現の有無. (3) 評価表現. 評価を表す語が出現する文は製品対比文の可能性が高い. (4) ルール A. ルールベースの手法におけるルール A(複数の製品名を含む) を満たすか否か. (5) 製品名の数 (0,1,2 以上). 製品名を多く含む文ほど製品対比文の可能性が高い. (6) 用言文節の数 (0,1,2 以上). 用言文節とは, ここでは用言(動詞, 形容詞など)を主辞とする文節とする. 用言文節を多く含む長い文ほど製品対比文の可能性が高い. (7) 評価表現と製品名が共に出現するか否か. (8) 製品名の係り先用言と (9) 製品名の間接的係り先用言. いずれも製品名と係り受け関係にある用言に特徴的な単語が現れるため, 素性とした.

提案手法の評価実験を行った. 対象製品として 11 個の製品を選び, 上述の手法でレビュー文の集合を得た. これらの文に対し, 製品対比文か否かを人手で判定し, 評価用データとした. 評価指数として, 正解率, 精度, 再現率及び F 値の 4 つを用いた. なお, 本論文で提案するルールベースの手法および機械学習に基づく手法における学習素性は,  $P_t=Nexus7$  のデータを精査し, 設計した. したがって, Nexus7 に対する評価はクロズドテストである.

まずルールベースの手法を評価した. ここではルール A, ルール B, 両者のいずれかの条件を満たすとき製品対比文と判定するルール A+B の 3 つを比較した. クロズドテストである  $P_t=Nexus7$  では, F 値の値はルール A+B(0.401), ルール A(0.236), ルール B(0.176) の順に高かった. ただし, それ以外の製品を対象にしたオープンテストでは, これらのルールの F 値の順序が製品カテゴリによって変化し, 一概にどのルールが有効であるかを判定することは難しい. ルール A について, 製品カテゴリによって F 値が大きく変化する理由として, 製品名の記述の仕方が製品カテゴリによって異なり, そのため製品名の検出が正しく行えなかったことがあげられる. 一方ルール B については, 製品カテゴリによってクロズドテストとオープンテストで大きな差が無かったことから, 様々な製品カテゴリに適用可能なルールとして対比表現辞書は有効であるといえる.

次に機械学習による手法を評価した. F 値は異なる製品カテゴリのレビュー文を訓練データとした場合 (0.171) と同じ製品カテゴリのレビュー文を 5 分割交差して訓練データとした場合 (0.369) で大きく差が出たことから, 製品対比文の特徴は製品カテゴリによって異なることがわかった. また, 同じ製品カテゴリのレビュー文を訓練データとしたときの SVM の F 値が 0.369 に対し, ルール A+B の F 値が 0.180 であることから, ルールベースの手法よりも機械学習に基づく手法の方が優れている. これは, 機械学習によって製品カテゴリの違いによる対比文の特徴をある程度自動的に学習できるためと考えられる.

全体的に, 製品対比文判定の正解率が高いが, 精度, 再現率, F 値は低かった. これは, 製品対比文の全体に占める割合が約 3% と低く, 製品対比文の正例がデータ内に多く出現しないことが原因と考えられる.

次に, 個々の学習素性の有効性を調べるために, 素性集合から素性を 1 つ除いて SVM を学習し, 全素性を用いて学習した SVM と比較した. この実験では, 訓練データは同じ製品カテゴリの文とし, 5 分割交差検定を行った. 製品対比文の判定に最も有効であった素性は自立語であったが, この素性は他の素性に比べて数が多いことから, 製品対比文の判定に最も大きく貢献することは自然である. F 値を基準にすれば, 自立語の次に有効な

のは、対比表現、ルール A であり、それ以外の素性については大差が無かった。これらの素性が自立語を除く他の素性と比べて有効であることは、ルールベースの手法で採用されたルールがある程度妥当であることを示唆する。同様に、製品カテゴリ毎に素性の有効性も調べた。自立語が一番有効であることは全ての製品で共通しているが、それ以外の素性については製品カテゴリによってその有効性にばらつきがあり、製品対比文判定の F 値の向上に貢献しない製品カテゴリも存在した。このことから、製品対比文をより正確に判定するためには今回提案したものとは異なる素性が必要である。

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 研究の背景	1
1.2 研究の目的	1
1.3 論文の構成	2
<b>第2章 関連研究</b>	<b>3</b>
2.1 評判情報の抽出に関する研究	3
2.2 製品間比較に関する研究	4
2.3 本研究の特色	5
<b>第3章 提案手法</b>	<b>6</b>
3.1 概要	6
3.2 前処理	6
3.2.1 レビュー文の取得	6
3.2.2 競合製品リスト	8
3.3 ルールベースの手法	10
3.3.1 ルール A-製品名の出現に基づく対比文の検出	10
3.3.2 ルール B-対比表現に基づく対比文の検出	11
3.4 機械学習に基づく手法	18
<b>第4章 評価実験</b>	<b>22</b>
4.1 データセット	22
4.2 実験結果	23
4.2.1 ルールベース手法の評価	24
4.2.2 機械学習に基づく手法の評価	27
4.3 製品対比文における製品名の出現の有無に関する考察	36
<b>第5章 結論</b>	<b>38</b>
5.1 まとめ	38
5.2 課題	38

# 第1章 序論

## 1.1 研究の背景

近年、個人が情報を発信する場としてウェブは一般的となっており、ブログやマイクロブログを通して膨大な数のユーザレビューが投稿されている。ユーザレビューとは、製品やサービスに対する使用経験、感想、批判などを述べた文である。また、近年ウェブ上で商品の注文、決済を行うECストアを通じた売買もまた広く普及し、実機に触れずに製品の購入を検討する機会が増えている。このため、他者による製品の感想や評価といったウェブ上に存在するユーザレビューの重要性が増している。ECストアなどが提供するユーザレビューの投稿場所では、投稿されたレビューからユーザが必要な情報を入手しやすいよう、閲覧ページを工夫しているものがある。例えばAmazon.co.jp<sup>1</sup>では、レビューのうち製品に対して肯定的なものと否定的なものを対比してユーザに提示している。価格.com<sup>2</sup>では、製品の持つ重量や機能などの属性について、それぞれ投稿されたレビューから該当する箇所を抽出・列挙しユーザに提示している。

しかしながら、ウェブ上に投稿されるレビュー文は非常に多く、またECストアのレビュー投稿サービスだけでなく、様々なブログ上にも散在していることから、必要としている情報や多様な知見を得ることが困難となっている。そのため、ユーザレビューを対象とした評判分析は、ユーザに対し有益な情報を提供する技術として注目されている。例えば、レビュー文が対象としている製品について肯定的であるか否定的であるかを自動的に判定し、その結果を提示することで製品に対する世間の評判を知ることができる [1]。

## 1.2 研究の目的

本研究では、製品レビューから製品に対する評判を抽出することを目的とする。この論文では、特に複数の製品を対比した文に着目する。例えば

iPad は Nexus7 と比べてサクサク動く

といった文は、iPad と Nexus7 を対比している。このような対比文は、タブレット端末の購入を検討しているユーザにとって、どの製品を購入するかを決定する指標として有益な情報であると考えられる。複数の製品に対する評判を取得するためには、個々の製品に対

---

<sup>1</sup><http://www.amazon.co.jp/>

<sup>2</sup><http://kakaku.com/>

してそのユーザレビューを検索し、その結果をまとめてユーザに提示する方法も考えられる。また、それぞれの製品の属性(例えば重量や価格)について評価している箇所を同定し、それをまとめることで製品の属性についての比較が行える。しかし、レビュー文から属性及びその評価を示す箇所を同定することはそれほど容易ではない。一方、対比文では“iPadは重い、Nexus7は軽い”のように重量という属性を比較しているが、文中から属性を検出する必要はなく、文を提示するだけで同じ属性(重量)を比較した文と理解できる。また“iPadは画面が綺麗だが、Nexus7は軽い”といった対比文でも、同じ属性ではないが暗にそれぞれの属性について“iPadは(他方に比べ)画面が綺麗だが、Nexus7は(他方に比べ)軽い”といった比較をしている。また、このような対比文では、1人のユーザが複数の製品を同時に評価しているという点で重要である。個々の製品毎にレビュー文を獲得し、同じ属性について言及する箇所を見つけて組み合わせるというアプローチは、製品毎に異なる人がレビュー文を書いているため、公正な比較になっているかに疑問が残る。例えば、「重い」「軽い」という感想は個人によって異なる。

本論文の目的は、競合するいくつかの製品の中から購入する製品を決めたいユーザに有益な情報を提供するために、複数の製品を対比した文(以下、製品対比文と呼ぶ)を検出する手法を提案することにある。ここでの製品対比文とは、単なる製品の仕様(価格、重量など)を比較した文だけではなく、ユーザの使用経験や感想を記した文も含むものとする。

### 1.3 論文の構成

本論文の構成は以下の通りである。2章では、関連研究としてウェブ上に存在する製品レビューからユーザにとって有用な情報を提示した手法について述べる。3章では、提案するシステムの概要を述べる。4章では、本研究で実装したシステムの評価について述べる。5章では、結論として本研究により得られた知見と今後の課題について述べる。

## 第2章 関連研究

本章では関連研究について述べる。本研究は、ユーザが購入を検討している複数の製品に関するレビュー文を収集し、それらを対比した製品対比文を検出することを目的としている。まず、製品に関する評判情報を抽出する研究について述べる。次に、複数の製品を比較した文を解析対象とした研究をいくつか挙げる。

### 2.1 評判情報の抽出に関する研究

乾・奥村は、テキストから評判情報を自動的に抽出する研究の動向を報告している [1]。彼らは、個人の評価を行っている情報を評価情報、評価情報が肯定的な評価であるか否定的な評価であるかを示す軸を評価極性、テキスト内で記述された評価情報の表現を評価表現、評価表現とその表現の持つ評価極性の組 (例えば (良い-肯定)) を集めたものを評価表現辞書と呼んでいる。これらは評判情報抽出の研究における重要な概念である。立石らは、製品に対する意見を < 対象, 属性, 評価値 > の組として定義し、評価を行う文書に特徴的に現れる評価語句を人手収集し作成した評価語辞書に基づくパターンマッチによって意見を抽出している [2]。小林らは、立石らの作成した特徴語辞書の作成を人手によるものからウェブ上のテキストから評価表現を半自動的に収集するよう拡張する方法を提案している [3]。彼女らはテキストから < 対象, 属性, 評価値 > のような典型的な共起パターンを利用し、共起パターンを満たす属性と評価値を表す表現を取得することで評価表現を収集している。

Xiaowen らは、評価表現の持つ感情極性が、製品カテゴリや評価を行う製品の特徴によって異なることに着目した [4]。例えば、評価表現「long」は、カメラのバッテリー寿命を評価する “the battery life is long” という文では肯定を示し、カメラのフォーカスが合うまでの時間を評価する “it takes long time to focus” という文では否定を示す。そして、製品の特徴ごとに各評価表現の持つ感情極性を自動的に判定する方法を提案している。具体的には、「great」や「disappointing」などの、製品カテゴリや製品の特徴に依存せず一定の感情極性を持つ評価表現の辞書と、「and」や「but」などの接続詞を基にした conjunction-rule やその拡張ルールを用いて判定を行っている。conjunction-rule とは、“this room is beautiful and spacious” のように二つの評価表現が and でつながれているとき、二つの評価表現が同じ感情極性を持つとするルールである [5]。Xiaowen らは、conjunction-rule の拡張ルールとして、“the camera battery has a long battery, which is great” のように明白な接続詞を持たないが「and」と同様の働きをする文章構造にも適用する Pseudo intra-sentence

conjunction rule や、「but」でつながれているときは二つの感情極性は異なるとした but clause rule などを提案した。そして製品の特徴毎にレビュー文を収集し、複数の評価表現が現れ、かつ「great」のように一定の感情極性を持つ評価表現も同時に現れているものを探し、conjunction-rule などを用いて評価表現の極性を自動的に判定した。

評判分析では、文書集合から文書、文、語句などの単位で、肯定あるいは否定の評価極性を判定するために評価表現辞書を利用することが多い。本研究においても、評価を行う文書に特徴的に現れる表現を検出するために評価表現辞書を利用している。

## 2.2 製品間比較に関する研究

Jindal と Liu は、テキストから2つの事物 (entity) 間の優劣関係を抽出する手法を提案している [8]。2つの事物を比較する文を自動検出し、それらの文から優劣を示す < 語句, 事物の属性, 事物 1, 事物 2 > からなる優劣関係を獲得する。具体的には、英語で事物の優劣関係を記述するとき、形容詞、副詞は比較級、最上級になるという特徴を基に、単語並びに品詞列のパターンを作成し、パターンマッチに基づく規則を用いて優劣関係を抽出している。例えば、以下の例文 1 からは <better, optics, canon, sony>, 例文 2 からは <higher, mileage, Car X, Car Y> という優劣関係を獲得する。

1. Canon's optics are better than those of Sony.
2. Car X has higher mileage per gallon than Car Y.

Ganapathibhotla らはこの手法を拡張し、単なる優劣関係を抽出するだけでなく、2つの事物を比較する文が事物 1, 事物 2 のいずれを優れていると評価しているかを判定する手法を提案している [7]。例えば、例文 1 のようにどの製品の特徴に対しても肯定の感情極性を示す「better」に対し、「mileage」の感情極性を求めるには製品カテゴリと製品の特徴に関する知識が必要となり、事物間の優劣を決定することは困難である。そこで Ganapathibhotla らは、レビュー文では一般的に冒頭に簡潔なまとめである Pros/Cons (図 2.1) が記述されていることに着目している。Pros/Cons にはそれぞれ製品の長所、短所が

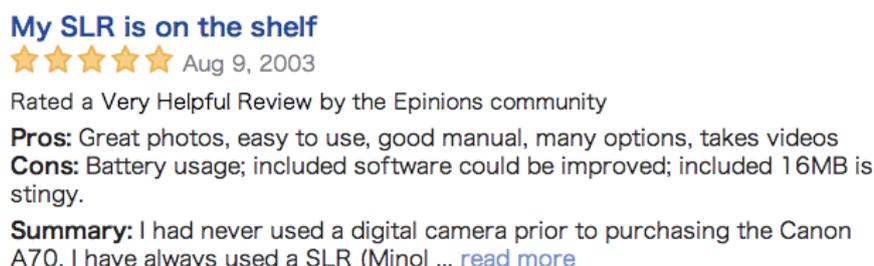


図 2.1: Pros/Cons を含むレビューの例

述べられており、製品の特徴と形容詞/副詞の組、ならびにその感情極性が明確に示されている。そこで、Pros/Cons のテキストを、「,」「.」「and」「but」で分割し製品の特徴のリストをまず作成する。図 2.1 の Pros からは以下のようなリストが得られる。左は Pros の行に出現する単語列、<> はそれから抽出した製品の特徴を表す。

great photos	<photos>
easy to use	<use>
good manual	<manual>
many options	<options>
takes videos	<video>

これらの特徴について、Pros に記述されていることから、それぞれ結合している形容詞/副詞は全て肯定の感情極性を示すと判定できる。逆に Cons から同様にして得た製品の特徴のリストについても、それぞれ結合している形容詞/副詞は全て否定の感情極性を示すと判定する。以上のようにして、製品の特徴に固有の評価語とその感情極性を取得することで、< 語句, 事物の属性, 事物 1, 事物 2 > からなる優劣関係において、どちらの事物が優れているかを決定する。

倉島らは比較表現と評価表現の出現から 2 製品間の優劣関係を求め、これらを集約して製品のランキングを作成する方法を提案している [6]。ここで比較表現とは、「より」「方が」などの何かを比較した文に特徴的に現れる表現としている。また、この研究は日本語を対象としている。

## 2.3 本研究の特色

2.2 節で述べた手法は主に事物の優劣関係を抽出することを目的しているのに対し、本論文では文が 2 つの事物を対比しているか否かを判定することに焦点を当て、ユーザによって与えられたクエリ (製品名) を含む製品レビューを分析し、他製品との対比を行っている文を抽出する。本研究では、“Nexus7 はステレオスピーカーを持ち、iPad は前面カメラを持つ”のように評価表現を持たないが、客観的な事実を対比した文も製品を比較している製品対比文として抽出を行う。対比文か否かの判定については、Jindal と Liu は、比較を示唆する 83 種類のキーワードを設定し、これらを素性とした分類器を Naive Bayes モデルならびに Support Vector Machine(SVM) で機械学習している [8]。一方、Xu らは単に 2 つの製品を含む文を対比文と判定している [9]。これに対し、本論文は、日本語の製品対比文に見られる特徴を分析し、この分析結果を基にした製品対比文の自動判定手法を提案する。また、製品対比文の判定手法が異なるタイプの製品を対象としたレビュー文に普遍的に適用できるかを実験的に検証する。

## 第3章 提案手法

### 3.1 概要

本論文が提案する手法の処理の流れを図3.1に示す。まず、ユーザが評判を知りたい対象製品 ( $P_t$  とする) についてレビュー文を取得する (3.2.1 項)。次に、対象製品  $P_t$  の競合製品リストを作成する (3.2.2 項)。次に、製品対比文の検出を行う。ここでは、与えられたレビュー文が製品対比文であるか否かを判定する2つの手法を提案する。一つはルールベースの手法 (3.3 項)、もう一つは機械学習に基づく手法 (3.4 項) である。これらの手法では文の形態素解析ならびに文節の係り受け解析を行う。本研究では形態素解析には MeCab<sup>1</sup>、文節の係り受け解析は CaboCha[10] を用いた。最後に、製品対比文であると判定した文を出力する。

### 3.2 前処理

#### 3.2.1 レビュー文の取得

ユーザが評判を知りたい対象製品  $P_t$  に対するレビュー文を取得する。 $P_t$  は “Nexus7” といった1つの製品名とする。

まず、“ $P_t$  レビュー” をクエリとして検索エンジンでウェブ検索を行う。検索されたウェブページからパッセージを取得する。パッセージとは、ここではHTMLのブロックタグで囲まれたテキストを指す。本研究で用いた21個のブロックタグの一覧を図3.2に示す。また、ブロックタグ以外のHTMLタグをインラインタグとして、パッセージの中に含まれているインラインタグを全て除去した。本研究で用いた49個のインラインタグの一覧を図3.3に示す。検索上位のウェブページから順に、パッセージを  $T$  個取得する。次に、パッセージを文に分割する。ここでは句点「。」をセパレータとしてテキストを文に分割した。以上の操作でレビュー文の集合を得る。もちろん、このようにして取得したレビュー文の中にはユーザが製品についての評価を行っていないものも含まれる。従って、ここで取得するのは正確にはレビュー文の候補である。これらの文は後続の処理で製品対比文であるか否かの判定を行うが、レビュー文でないものは製品対比文ではないと判定する。

---

<sup>1</sup><http://code.google.com/p/mecab/>

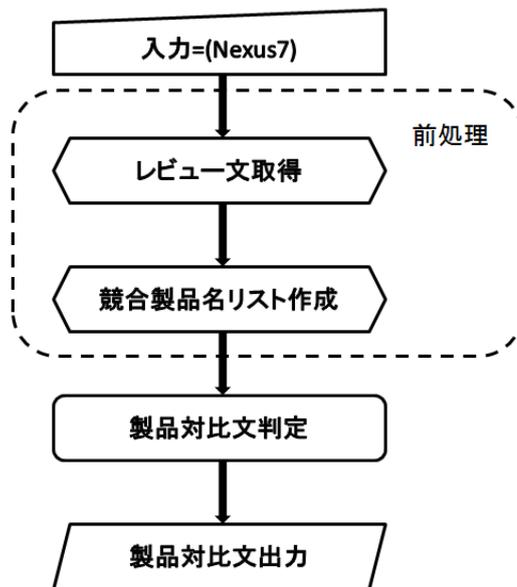


図 3.1: システムの概要

address, blockquote, center, div, dl, fieldset, form, h1, h2, h3, h4, h5, h6,  
hr, noframes, noscript, ol, p, pre, table, ul

図 3.2: ブロックタグの一覧

br, tt, i, b, u, s, strike, big, small, em, strong, dfn, code, samp, kbd, var,  
cite, abbr, acronym, a, img, applet, object, font, basefont, map, q, sub, sup,  
span, bdo, iframe, input, select, textarea, label, button, blink, embed, layer,  
ilayer, nolayer, nobr, wbr, ruby, rb, rp, rt, spacer

図 3.3: インラインタグの一覧

### 3.2.2 競合製品リスト

提案手法では，製品対比文判定の際に，対象製品  $P_t$  と競合する製品のリストを利用する．競合製品とは， $P_t$  と同じ種類の製品で，購入を検討する際に比較の対象となりうる製品とする．本論文では，競合製品リストは以下の手続きで半自動的に獲得する．

- ウェブサイト「価格.com<sup>2</sup>」において  $P_t$  が属する製品カテゴリを求める．
- 価格.com における製品カテゴリのスペック検索のページを表示させる．ただし，検索キーワードは指定しない．
- そのページに登録されている全ての製品のリストを (編集せずに) そのまま競合製品リストとする．

以下，競合製品リストに登録された製品を  $P_c$  とする．例えば  $P_t = \text{Nexus7}$  とした場合，「価格.com」で Nexus7 の製品カテゴリである“タブレット PC(端末)・PDA”のスペック検索のページを表示させる．このページのスナップショットを図 3.4 に示す．図 3.4 の下半分には，製品カテゴリ“タブレット PC(端末)・PDA”に属する製品の一覧が表示されている．この表から iPad, Kindle, Surface などの製品名を取得し， $P_c$  として競合製品リストに加える．

後述する手法では，製品名  $P_t$  や  $P_c$  が文中に出現するか否かを判定する．この際，製品名が完全に一致する場合だけでなく，文中の単語列と製品名を構成する単語列が先頭から部分的に一致するときも製品名が出現すると判定する．例えば， $P_c$  が「Reader Wi-Fi モデル PRS-T1」のとき，文中に「Reader」という単語があれば，この製品が文中に出現すると判定する．

---

<sup>2</sup><http://kakaku.com/>

タブレットPC(端末)・PDA スペック検索

ホーム > パソコン > タブレットPC(端末)・PDA > タブレットPC(端末)・PDA スペック検索

メーカー名	選択しない	
シリーズ名	選択しない	
基本スペック	タイプ	コア数
	CPU周波数	メモリー
OS	記憶容量	詳細OS種類
	OS種類	
液晶	液晶サイズ	ppi
	画面解像度	
バッテリー	バッテリー種類	バッテリー容量
カードスロット	<input type="checkbox"/> SDカード	<input type="checkbox"/> microSDカード
	<input type="checkbox"/> microSDHCカード	<input type="checkbox"/> コンパクトフラッシュ
インターフェース	<input type="checkbox"/> microUSB	<input type="checkbox"/> miniUSB
	<input type="checkbox"/> USB	<input type="checkbox"/> HDMI端子
センサー	<input type="checkbox"/> 加速度センサー	<input type="checkbox"/> ジャイロセンサー
	<input type="checkbox"/> 光センサー	<input type="checkbox"/> デジタルコンパス
通信機能	<input type="checkbox"/> IEEE802.11b	<input type="checkbox"/> IEEE802.11g
	<input type="checkbox"/> IEEE802.11n	<input type="checkbox"/> Bluetooth
カメラ	<input type="checkbox"/> ワンセグ	<input type="checkbox"/> 地上デジタル
	<input type="checkbox"/> クレドル	<input type="checkbox"/> キーボード
その他機能	<input type="checkbox"/> HDD搭載	<input type="checkbox"/> Google Play対応
	<input type="checkbox"/> 防水・防滴性能	<input type="checkbox"/> 統合ソフト
サイズ・重量	幅	高さ
	奥行	重量
カラー	カラー	
発売時期	選択しない	
価格	円以下 <input type="checkbox"/> 価格がある製品のみ絞る	
画像	<input type="checkbox"/> 画像を表示しない	
並び順	クチコミの多い順	

↑ 選択した項目( )が、検索結果に表示されます。

検索

896件中 1-30件の製品が表示されています

表示形式:  色をまとめて表示  色別に表示

チェックをいれてクリック 詳しく比較する	製品名	価格	取扱店舗数	ランキング	満足度	クチコミ	基本スペック				液晶						
							タイプ	OS種類	CPU	コア数	メモリー	記憶容量	センサー	液晶サイズ	液晶タイプ	画面解像度	ppi
<input type="checkbox"/>	Google Nexus 7 Wi-Fi+モバイル通信対応モデル 32GB [2012]	¥21,800	22店舗	36位	4.56 (36票)	20003件	タブレット	Android 4.2	Tegra 3 1.3GHz	4	1GB	32GB	加速度センサー ジャイロセンサー 光センサー デジタルコンパス	7インチ	IPS	1280x800	216
<input type="checkbox"/>	Google Nexus 7 Wi-Fiモデル 16GB [2012]	¥16,279	6店舗	61位	4.53 (382票)	20003件	タブレット	Android 4.1	Tegra 3 1.3GHz	4	1GB	16GB	加速度センサー ジャイロセンサー 光センサー デジタルコンパス	7インチ	IPS	1280x800	216
<input type="checkbox"/>	Google Nexus 7 Wi-Fiモデル 32GB [2012]	¥16,180	28店舗	22位	4.58 (219票)	20003件	タブレット	Android 4.1	Tegra 3 1.3GHz	4	1GB	32GB	加速度センサー ジャイロセンサー 光センサー デジタルコンパス	7インチ	IPS	1280x800	216
<input type="checkbox"/>	APPLE iPad Wi-Fiモデル 16GB MB292J/A	-	0店舗	-	4.36 (91票)	10019件	タブレット	iOS						9.7インチ	IPS	1024x768	
<input type="checkbox"/>	APPLE iPad Wi-Fiモデル 32GB MB293J/A	-	0店舗	-	4.72 (57票)	10019件	タブレット	iOS						9.7インチ	IPS	1024x768	

図 3.4: ウェブサイト「価格.com」における Nexus7 を含む製品カテゴリの製品一覧

## 3.3 ルールベースの手法

### 3.3.1 ルール A-製品名の出現に基づく対比文の検出

ルール A は製品名が複数出現する文を製品対比文と判定する規則である。これは、製品名が2つ以上現れる文においては、それらの製品が対比されているという考えに基づく。具体的には、図 3.5 に示す2つの条件を満たすとき、製品対比文として判定する。

条件 1. 複数の製品が文内に出現する

条件 2. 製品名を含む節が2つ以上存在する

図 3.5: ルール A

図 3.5 の条件 1 では、まず対比文は製品名を複数含まなければならないとした。ただし、製品名が複数現れた場合でも、出現した製品を同列に扱って評価を行っているときなど、それらの製品を対比しているとはみなされない場合がある。例えば

iPad も Nexus7 も十分高性能

というレビュー文には製品名が複数現れているが、iPad と Nexus7 の両方についての評価であり、両者を対比しているわけではない。このようなレビュー文は製品対比文ではないと判定したい。そのため図 3.5 の条件 2 を設定した。条件 2 では、レビュー文を節に分割し、2つ以上の節に異なる製品名が出現しなければならないとした。ここでは、「節」とは用言文節もしくは接続詞文節で区切られた文字列と定義する。用言文節とは、主辞の品詞が動詞、形容詞である文節、もしくは名詞+「だ」で構成される名詞述語となっている文節を指す。一方、接続詞文節は、主辞の品詞が接続詞である文節を指す。

ルール A による判定例を以下に示す。

- タブレット A は画面が 大きいが| タブレット B は 小さい|  
⇒ ○
- タブレット A もタブレット B も性能は いいが| 持ち歩くには 大きすぎる|  
⇒ ×

これらの例において、下線は用言文節を、|は節の区切りを表す。最初の文は2つの節から構成され、それぞれの節に「タブレット A」、「タブレット B」という製品名が存在する。したがってルール A の条件 1, 2 を満たすため、製品対比文と判定する。次に、2番目の文について考察する。この文も2つの節で構成されている。「タブレット A」、「タブレット B」という2つの製品名が出現し、条件 1 を満たす。しかし、これらは同じ節に出現し、製品名を含む節は1つしかない。従って条件 2 を満たさないため、製品対比文ではないと判定する。

### 3.3.2 ルール B-対比表現に基づく対比文の検出

ルール B では人手作成した対比表現辞書を基に対比文の判定を行う。対比表現辞書とは、対比文に特徴的に現れる表現を対比表現とし、これを収集した辞書である。対比表現にはそれぞれ付加条件を設定し、文内で対比表現が付加条件を満たして現れている場合に、その文を製品対比文として判定する。

以下、対比表現辞書における対比表現と付加条件の詳細を述べる。

- **だ**

助動詞「だ」の假定形「なら」が文中に出現するとき、製品対比文と判定する。この対比表現を含む製品対比文の例を示す。

1. iPadならどの端末も iOS を最新バージョンにアップデートできる。
2. 寝っ転がってネットサーフィンやゲームアプリをする ならnexus7.

付加条件

- 活用形が假定形である  
つまり、文中における出現形が「なら」である。
- 「なら」と同じ文節内に製品名を含む  
「なら」の前に製品名ではない名詞が出現したときは、製品を対比しているわけではない。そのため、例文 1 のように「なら」を含む文節に製品名が出現しなければならないという付加条件を加えた。
- 「なら」の直前に動詞が現れる。ただし動詞と「なら」の間に「たい」「の」という単語は例外として出現してもよい。  
例文 2 のように「(動詞)なら」という表現があったとき、製品対比文と判定するための条件である。ただし、製品対比文の中には「ゲームアプリをするなら」の他に「ゲームアプリをするのなら」、「ゲームアプリをしたいのなら」という表現も見られたため、例外として「たい」「の」という単語は動詞と「なら」の間に出現してもよいとした。

- **でも、ならでは**

「でも」「ならでは」という表現が文中に出現するとき製品対比文と判定する。この対比表現を含む製品対比文の例を以下に示す。

1. iPad で bluetooth 接続で使っていますが、Nexus7でも使える。
2. Nexus7ならではのカスタマイズ性。

付加条件

- 対比表現と同じ文節内に製品名を含む  
「でも」「ならでは」の前に製品名ではない名詞が出現しているときは製品を対比しているわけではないため、このような付加条件を設定した。

- より

「より」という表現が文中に出現するとき製品対比文と判定する。この対比表現を含む製品対比文の例を以下に示す。

1. 圧倒的に Nexus7の方が iPadより 使いやすいです。
2. Nexus7で写真を撮る より きれい。

付加条件

- 「より」と同じ文節内に製品名を含む  
「より」の前に製品名ではない名詞が出現しているときは製品を対比しているわけではないため、このような付加条件を設定した。例文1はこの付加条件を満たす。
- 「より」と同じ文節内に動詞を含む  
「より」の前に製品名が出現しなくても、例文2のように「(動詞)より」という表現があれば製品を対比している可能性があるため、この付加条件を設定した。

- しか

「しか」という表現が文中に出現するとき製品対比文として判定する。この対比表現を含む製品対比文の例を以下に示す。

1. Nexus7に しか ないアプリもある。

付加条件

- 「しか」の直前に製品名が出現する。ただし製品名と「しか」の間に「に」という単語は例外として出現してもよい。  
「しか」の前に製品名ではない名詞が出現しているときは製品を対比しているわけではない。そのため「しか」の直前は製品名でなければならないという条件を加えた。ただし、例文1のように、製品対比文の中には「Nexus7にしか」という表現も見られたため、例外として「に」という単語は出現してもよいとした。

- 比較

「～と比較すると」などの表現があるとき製品対比文として判定する。この対比表現を含む製品対比文の例を以下に示す。

1. 機能や性能を 比較 すると Nexus7のコストパフォーマンスはとても高いです。

2. これも iPad mini との 比較 ですが、もうちょっとバッテリーがもって欲しいかなーと思います

#### 付加条件

- 用言文節に出現し、同じ文節内に「べき」「たら」が出現せず、かつレビュー文内に用言文節が複数出現する。

「比較」はサ変名詞なので、文中では名詞として使われることもあれば、動詞として使われることもある。用言文節に出現するという条件は「比較」が動詞と使われていると言うことを表す。レビュー文内に用言文節が複数出現するという条件は、用言文節を複数含むような文でないという製品対比文にはならない、すなわち製品を対比している文は複数の節を含むような長い文であるという仮定に基づいている。例えば以下の例文 a では「比較」が出現しているが、1つの節からなる単純な文であり、製品対比文ではない。また、例文 b のように、「比較」が出現している、その後「べき」「たら」も出現した場合は対比文ではないため、製品対比文と判定しないよう条件を設けた。

- a. 価格を比較した。
- b. 近い価格帯の製品と性能を比較するべき。

なお、上記の例文 1 はこの付加条件を満たす。

- 製品名が「比較」に係り、かつ「比較」が文末以外に出現する。  
「比較」が名詞として使われているときでも、製品名から係られている場合は製品の対比が行われる。例えば、上記の例文 2 では「iPad mini との」という文節が「比較ですけど、」という文節に係り、iPad mini と他の製品が対比されていることがわかる。この条件は文節の係り受け解析の結果を見てチェックする。したがって、ここでの条件は正確には、製品名を含む文節が「比較」を含む文節に係り、かつ「比較」を含む文節が文末に出現しない、となる。ただし、例文 a のように、「比較」が文末に現れる場合、その文では対比が行われていないことが多いため、「比較」が出現する文節は文末ではないという条件を設けた。

- a. iPad と Nexus7 の価格比較

#### ● 方

「方」という表現が文中に出現するとき製品対比文として判定する。この対比表現を含む製品対比文の例を以下に示す。

1. 圧倒的に Nexus7 の 方 が iPad より使いやすいです。

#### 付加条件

- 「方」の読みが「ホウ」である  
「方」が「カタ」という読みの際は対比表現とは見なさない。
- 「方」の後ろに「が」が続く  
以下の例文 a のように、「方」が出現する文は常に製品対比文になるというわけではない。「方が」と言う表現があるときのみ製品対比文となる。この付加条件を加えることにより、例文 a のような文を製品対比文として誤検出することを妨げることができる。

a. 店頭の方にはある

- 敵う, 劣る, 優れる, 比べる, 較べる, 勝つ, 負ける  
「劣る」などの表現が文中に出現するとき製品対比文と判定する。この対比表現を含む製品対比文の例を以下に示す。

1. 処理速度は Nexus に 劣る, 負荷がかかるものには適しません。

付加条件

- 文内に用言文節が複数出現する  
これは用言文節を複数含むような長い文でないという製品対比文にはならないという仮定に基づく。

- 比較的, 圧倒的, 同等, 同じ, 勝利, 圧勝  
「比較的」などの表現が文中に出現するとき製品対比文と判定する。この対比表現を含む製品対比文の例を以下に示す。

1. iPad mini は 比較的 軽量だが幅広で, 実質的に片手で持てない。

付加条件

- なし  
これらの対比表現が出現したときは無条件で製品対比文と判定する。

対比表現辞書では、いくつかの対比表現の付加条件として「文中に用言文節が複数含まれる」という条件を設定している。前述のように、これは用言を複数含むような長い文でないという製品対比文にはならないという仮定に基づく。しかしこの仮定は常に正しいわけではない。「iPad は Nexus7 より画質が優れている」という文のように、用言が1つしかない文でも対比文になることがある。この場合、対比表現「優れる」については付加条件を満たさない。しかし、短い対比文には別の対比表現が存在する可能性が高い。この例では「より」という別の対比表現があるため製品対比文と判定される。このとき「より」は対比を強く示唆する表現であるため、用言文節を複数含むという付加条件は設定していないことに注意されたい。

上記の 21 個の対比表現とその付加条件を表 3.1, 3.2 にまとめる。この表では、対比表現の基本形と共に、それらの対比表現が満たすべき付加条件を、品詞、活用形、読み、その他に分けて記述している。対比表現の中には複数の付加条件が示されているものがあり、表 3.1, 3.2 で同じ行に存在する条件は全て満たさなければならない。また、「だ」「より」「比較」は同じ対比表現が点線で 2 行に分かれているが、この場合はいずれかの条件を満たせば製品対比文と判定する。

ルール B は以下のように適用される。まず、文を形態素解析し、それぞれの単語の基本形が表 3.1, 3.2 の対比表現にマッチするかを調べる。対比表現が見つかったとき、その付加条件を満たすかをチェックする。付加条件が満たされたとき、その文を製品対比文と判定する。

表 3.1: 対比表現と対比文検出のための付加条件 (その 1)

対比表現 (基本形)	条件 (品詞)	条件 (活用形)	条件 (読み)	条件 (その他)
だ	助動詞	仮定形		同じ文節内に製品名を含む
だ	助動詞	仮定形		同じ文節内に動詞を含む 動詞と対比表現の間に“たい”, “の”以外の単語が出現しない
でも				同じ文節内に製品名を含む
ならでは				同じ文節内に製品名を含む
より				同じ文節内に製品名を含む
より				同じ文節内に動詞を含む
しか	助詞			同じ文節内に製品名を含む 対比表現の前に“に”以外の単語が出現しない
比較				用言文節に出現する 文内に用言文節が複数出現する
比較				同文節内に“べき”, “たら”が出現しない 製品名が「比較」に係るの 文末以外に出現する
方			ホウ	対比表現の後ろに“か”が続く

表 3.2: 対比表現と対比文検出のための付加条件 (その 2)

対比表現 (基本形)	条件 (品詞)	条件 (活用形)	条件 (読み)	条件 (その他)
比較的				
圧倒的				
同等				
同じ				
勝利				
圧勝				
敵う	動詞			文内に用言文節が複数出現する
勝る	動詞			文内に用言文節が複数出現する
劣る	動詞			文内に用言文節が複数出現する
優れる	動詞			文内に用言文節が複数出現する
比べる	動詞			文内に用言文節が複数出現する
較べる	動詞			文内に用言文節が複数出現する
勝つ	動詞			文内に用言文節が複数出現する
負ける	動詞			文内に用言文節が複数出現する

### 3.4 機械学習に基づく手法

対比文がタグ付けされた製品レビューの集合を訓練データとし、与えられた文が製品対比文か否かを判定する二値分類器を教師あり学習する。本論文で採用した学習アルゴリズムは Support Vector Machine(SVM) である。SVM の学習には libSVM<sup>3</sup> を用いた。カーネルは線形カーネルを用い、それ以外のパラメータは libSVM のデフォルト値とした。

以下、SVM の学習に用いた素性について説明する。なお、以下の説明において、()内は素性を表す略記号で、本論文の以降の説明で用いる。

- 文中の自立語 (bow)  
レビュー文に出現する自立語の基本形を素性とする。いわゆる bag of words と呼ばれる素性である。例えば、以下の文では太字で表された単語が自立語である。

Apple の**商品**と**比べる**と**見劣り**するけど**十分合格点**

したがって、自立語の素性は以下ようになる。ただし、「bow;」は素性の種類を表す識別子である。

bow;apple bow;商品 bow;比べる bow;見劣り bow;十分 bow;合格 bow;点

- 対比表現 (comExp)  
3.3.2 項で作成した対比表現辞書内の対比表現を素性とする。ただし、3.3.2 項で規定した付加条件も満たさなければならない。ルールベースの手法で用いた対比表現辞書が製品対比文の判定に有効であるとして設計した。例えば、以下の例文では太字の対比表現があり、かつ付加条件も満たす。

Apple の商品と**比べる**と**見劣り**するけど**十分合格点**

したがって、以下の素性が抽出される。

comExp; 比べる

- 評価表現 (evalExp)  
レビュー文内に評価表現が含まれるかを 0, 1 の二値で表す素性である。レビュー文内に「きれい」などの評価表現があれば、製品についての評価を述べている文の可能性が高い。ここでは日本語評価極性辞書(用言編)[3]を用い、辞書の中の評価表現が文中に出現するときに 1, それ以外は 0 とする。評価表現があるだけでは製品対比文とは限らないが、製品名の有無の素性などと組み合わせることで製品対比文の判定に有効になると考えた。例えば、以下の例文では太字の評価表現がある。

その共通のアプリで特に Android の方が使い勝手が**いい**ソフトが「Gmail」です

<sup>3</sup><http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

したがって、この素性の値は以下のように1となる。

evalExp;1

- ルール A(ruleA)

3.3.1 項で述べたルール A の条件を満たすか否かを 0, 1 の二値で表す素性である。ルール A は文中に複数の製品名が出現するかをチェックするルールであるが、ルール A の条件を満たす文は製品対比文である可能性が高い。例えば、以下の例文はルール A の条件を満たす。(| は節の境界を表す)

iPad は画面が大きい | が | Nexus7 は小さい |

したがって、この素性の値は以下のように1となる。

ruleA;1

- 製品名数 (num\_P)

レビュー文内に出現する製品名の数を 0 個, 1 個, 2 個以上の 3 値で表す素性である。製品名が多く出現する文ほど製品対比文である可能性が高いと考え、この素性を設定した。例えば、以下の例文では太字で示された製品名が 2 つ出現する。

iPad は画面が大きい | が | **Nexus7** は小さい

したがって、この素性の値は以下のように2となる。

num\_P;2

- 用言文節数 (num\_Y)

レビュー文内に出現する用言文節の数を 0 個, 1 個, 2 個以上の 3 値で表す素性である。長い文ほど製品対比文になりやすいと考え、文の長さを用言文節の数で測るための素性である。文の長さは文字数や単語数でも測ることができるが、素性の数が多くなり、過学習を起ししやすい。このため、本研究では用言文節の数で文の長さを測った。例えば、以下の例文では下線をつけた 2 つの用言文節がある。

iPad は画面が 大きい | が | Nexus7 は 小さい

したがって、この素性の値は以下のように2となる。

num\_Y;2

- 評価表現&製品名 (evalExp&P)

評価表現と製品名が同時にレビュー文内に出現するか否かを 0, 1 の二値で表す素性である。これは評価表現と製品名の素性の組み合わせである。例えば、以下の例文では「iPad」という製品名と「重い」という評価表現がある。

iPadの方が重い

したがって、この素性の値は以下のように1となる。

evalExp&P;1

- 製品名の係り先用言 (Y-P\_direct)

製品名を含む文節の係り先が用言文節の時、係り先文節の主辞の基本形を素性とする。製品名を含む文節が直接係る用言節の主辞に製品対比文の特徴が現れるのではないかと考え設計した。例えば、以下の例文では「iPadは」という文節が「大きい」に係る。

iPadは画面のサイズが大きい

したがって、以下の素性が得られる。

Y-P\_direct:大きい

- 製品名の間接的係り先用言 (Y-P\_indirect)

製品名を含む文節の間接的な係り先が用言文節の時、係り先文節の主辞の基本形を素性とする。Y-P\_directと類似しているが、直接の係り受け関係ではなく、文節の係り受け関係を辿って到達する間接的な係り受けを考慮した素性である。例えば、以下の例文は「iPadの」という文節が「方が」に係り、「方が」という文節が「大きい」に係る。

iPadの方が画面が大きい

したがって、以下の素性が得られる。

Y-P\_indirect:大きい

以上の学習素性を表3.3にまとめる。

表 3.3: 学習素性の一覧

素性	略称	説明
文中の自立語	bow	文中に出現する自立語の基本形
対比表現	comExp	文中に出現する対比表現
評価表現	evalExp	文中に評価表現が出現するか否か
ルール A	ruleA	3.3.1 項で述べたルール A を満たすか否か
製品名数	num_P	文中に出現する製品名の数
用言節数	num_Y	文中に出現する用言節の数
評価表現&製品名	evalExp&P	評価表現と製品名が同一文節内に出現するか否か
製品名の係り先用言	P-Y_direct	製品名を含む文節の係り先が用言節の時, 係り先用言の主辞
製品名の間接的係り先用言	P-Y_indirect	製品名を含む文節の間接的な係り先が用言の時, その係り先用言の主辞

## 第4章 評価実験

本章では、提案手法の評価実験について述べる。

### 4.1 データセット

製品対比文の検出を行う対象製品として、表 4.1 に示す  $P_{t1} \sim P_{t11}$  の 11 個の製品を選んだ。これらの対象製品はそれぞれ製品カテゴリは異なる。 $P_{t11} = \text{CPU}$  は例外的に製品名ではなく製品カテゴリを対象製品とした。これは、製品の購入を検討する際、特定の候補とする製品がないときには製品カテゴリをクエリとしてレビュー文を検索する場合を想定している。次に、対象製品毎に 3.2.1 項で述べた手続きで、その製品のレビュー文を取得した。「 $P_t$  レビュー」をクエリとし、検索エンジンを用いて製品レビューを含むと思われる文書を収集し、上位の文書から  $T$  個のパスセージを取得した。ここでは  $T=2000$  とした。本実験では検索エンジンとして Google<sup>1</sup> を用いた。さらに、パスセージを文に分割し、レビュー文の集合を得た。次に、こうして得られた個々の文に対して製品対比文であるか否かのタグ付けを行った。

それぞれの対象製品について取得した上位のパスセージ数  $T$  とレビュー文数、ならびに製品対比文と判定した文の数を表 4.2 に示す。ここでは、対象製品  $P_t$  が出現しない場合も、対象製品の属性について対比を行っている文は製品対比文と判定した。例えば、対象製品名が  $\text{Nexus7}(P_{t1})$  のとき、“Android では～だが iOS では～だ” といったようにタブレットの OS を比較している文は、対象製品が出現しなくても製品対比文とした。 $P_{t1}$  については後述の理由で  $T=4000$  とした。また  $P_{t4}$  は十分な量の文書が検索できなかったため  $T=700$  とした。

本論文で提案したルールベースの手法、及び機械学習に基づく手法における学習素性は  $P_{t1} = \text{Nexus7}$  のデータを精査し、設計した。従って、 $P_{t1}$  を対象とした実験はクローズドテストである。また、より多くのレビュー文を調べるために、 $P_{t1}$  については取得するパスセージの数  $T$  を 4000 とした。

---

<sup>1</sup><http://www.google.co.jp/>

表 4.1: 製品一覧

	対象製品	製品カテゴリ
$P_{t1}^*$	Nexus7	タブレット PC(端末)・PDA
$P_{t2}$	REGZA 26ZP2 [26 インチ]	液晶テレビ
$P_{t3}$	うるおい 光クリエール ACK70N	空気清浄機
$P_{t4}$	AVR-X4000	AV アンプ
$P_{t5}$	ブルーレイディーガ DMR-BWT650	ブルーレイ・DVD レコーダー
$P_{t6}$	EH-TW5200	プロジェクタ
$P_{t7}$	ラムダッシュ ES-ST25	シェーバー
$P_{t8}$	MDR-DS7500	ヘッドホン・イヤホン
$P_{t9}$	極め炊き NP-BA10	炊飯器
$P_{t10}$	SD1 Merrill	デジタル一眼カメラ
$P_{t11}$		CPU

表 4.2: 実験に用いたレビュー文, 並びに製品対比文の数

対象製品	T	文	対比文
$P_{t1}^*$ Nexus7(タブレット)	4000	6535	314
$P_{t2}$ REGZA(液晶テレビ)	2000	3084	165
$P_{t3}$ うるおい 光クリエール (空気清浄機)	2000	2922	56
$P_{t4}$ AVR-X4000(AV アンプ)	2000	3170	24
$P_{t5}$ ブルーレイディーガ (レコーダー)	2000	2658	24
$P_{t6}$ EH-TW5200(プロジェクタ)	2000	3027	57
$P_{t7}$ ラムダッシュ(シェーバー)	2000	3138	122
$P_{t8}$ MDR-DS7500(ヘッドホン)	2000	3905	135
$P_{t9}$ 極め炊き (炊飯器)	700	941	6
$P_{t10}$ SD1 Merrill(デジカメ)	2000	3540	122
$P_{t11}$ (CPU)	2000	2712	171

## 4.2 実験結果

3.3 節及び 3.4 節で述べた手法による製品対比文判定の評価を行う。評価指標として、以下の 4 つを用いた。

- 正解率 (A)

システムの判定結果とレビュー文にタグ付けされた対比文か否かを示すタグが一致している割合。定義を式 (4.1) に示す。

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

- 精度 (P)

システムによって製品対比文であると判定された文のうち，その判定が正しい文の割合．定義を式 (4.2) に示す．

$$P = \frac{TP}{TP + FP} \quad (4.2)$$

- 再現率 (R)

製品対比文であるとタグ付けされたレビュー文のうち，システムによって製品対比文として検出できた文の割合．定義を式 (4.3) に示す．

$$R = \frac{TP}{TP + FN} \quad (4.3)$$

- F 値

精度と再現性の調和平均．定義を式 (4.4) に示す．

$$F = \frac{2 \cdot P \cdot R}{P + R} \quad (4.4)$$

式 (4.1), (4.2), (4.3), (4.4) における記号について説明する．まず，タグ付けされたレビュー文の集合のうち，製品対比文であるとタグ付けが行われたレビュー文を正例，製品対比文ではないとタグ付けが行われたレビュー文を負例とする． $TP$  は正例に対してシステムが製品対比文であると判定した文の数， $TN$  は負例に対しシステムが製品対比文でないと判定した文の数， $FP$  は負例に対しシステムが製品対比文であると誤って判定した文の数， $FN$  は正例に対しシステムが製品対比文ではないと誤って判定した文の数である．今回の実験では 11 個の製品を対象としているため，個々の製品に対する製品対比文検出の評価指数と共に，その平均も示す．以降に示す実験結果の表では，11 個の製品に対する評価指数のマクロ平均を「平均 (all)」，クローズドテストである  $P_{t1}$  を除いた 10 個の製品カテゴリに対するマクロ平均を「平均」の行で示す．

#### 4.2.1 ルールベース手法の評価

3.3 節で述べたルールベースの手法を評価する．ルール A で判定を行った場合の実験結果を表 4.3，ルール B で判定を行った場合の実験結果を表 4.4，ルール A もしくはルール B のいずれかを満たす場合を製品対比文であるとして判定を行った場合の実験結果を表 4.5 に示す．

提案手法 (ルール A+B) の F 値は 0.18 であった．クローズドテストである  $P_{t1}$  では比較的 F 値が高い (0.401) が，それ以外の製品では F 値は大きく劣る．ルール A 及びルール B は Nexus7 のレビュー文を基に設計したが，他のカテゴリの製品については必ずしも有効ではないことがわかる．また，製品毎の F 値のばらつきも大きい．したがって，表 4.5 の実験結果は，製品対比文の特徴は製品カテゴリによって異なることを示唆する．

表 4.3: 実験結果 (ルールベース手法; ルール A)

対象製品	A	P	R	F
$P_{t1}^*$ Nexus7(タブレット)	.952	.522	.152	.236
$P_{t2}$ REGZA(液晶テレビ)	.945	.167	.006	.012
$P_{t3}$ うるおい 光クリエール(空気清浄機)	.946	-	-	-
$P_{t4}$ AVR-X4000(AV アンプ)	.992	.400	.167	.235
$P_{t5}$ ブルーレイディーガ(レコーダー)	.988	-	-	-
$P_{t6}$ EH-TW5200(プロジェクタ)	.980	.200	.018	.032
$P_{t7}$ ラムダッシュ(シェーバー)	.951	.119	.041	.061
$P_{t8}$ MDR-DS7500(ヘッドホン)	.965	.421	.059	.104
$P_{t9}$ 極め炊き(炊飯器)	.994	-	-	-
$P_{t10}$ SD1 Merrill(デジカメ)	.965	-	-	-
$P_{t11}$ (CPU)	.916	.293	.240	.264
平均	.964	.267	.089	.118
平均 (all)	.963	.303	.098	.135

表 4.4: 実験結果 (ルールベース手法; ルール B)

対象製品	A	P	R	F
$P_{t1}^*$ Nexus7(タブレット)	.940	.139	.238	.176
$P_{t2}$ REGZA(液晶テレビ)	.933	.312	.212	.253
$P_{t3}$ うるおい 光クリエール(空気清浄機)	.953	.113	.214	.148
$P_{t4}$ AVR-X4000(AV アンプ)	.954	.023	.125	.039
$P_{t5}$ ブルーレイディーガ(レコーダー)	.977	.140	.292	.289
$P_{t6}$ EH-TW5200(プロジェクタ)	.964	.019	.018	.018
$P_{t7}$ ラムダッシュ(シェーバー)	.946	.236	.172	.199
$P_{t8}$ MDR-DS7500(ヘッドホン)	.945	.232	.259	.245
$P_{t9}$ 極め炊き(炊飯器)	.978	.105	.333	.160
$P_{t10}$ SD1 Merrill(デジカメ)	.923	.139	.238	.176
$P_{t11}$ (CPU)	.912	.343	.433	.382
平均	.949	.166	.230	.191
平均 (all)	.948	.164	.230	.190

次に、ルール A, ルール B, ルール A+B の比較を行う。  $P_{t1}$  を対象としたクローズドテストでは、F 値の値は、ルール A+B(0.401), ルール A(0.236), ルール B(0.176) の順に高い。これは、Nexus7 を対象製品としたとき、複数の製品名を含むかを判定するルール Aの方が対比表現を含むかを判定するルール B よりも有効であることがわかる。また、ルール A とルール B を併用することで、製品対比文の判定の F 値がかなり向上することもわ

表 4.5: 実験結果 (ルールベース手法; ルール A+ルール B)

対象製品	A	P	R	F
$P_{t1}^*$ Nexus7(タブレット)	.938	.374	.432	.401
$P_{t2}$ REGZA(液晶テレビ)	.931	.299	.212	.248
$P_{t3}$ うるおい 光クリエール(空気清浄機)	.921	.061	.214	.094
$P_{t4}$ AVR-X4000(AV アンプ)	.954	.051	.292	.088
$P_{t5}$ ブルーレイディーガ(レコーダー)	.975	.123	.292	.173
$P_{t6}$ EH-TW5200(プロジェクタ)	.963	.034	.035	.035
$P_{t7}$ ラムダッシュ(シェーバー)	.936	.179	.180	.180
$P_{t8}$ MDR-DS7500(ヘッドホン)	.944	.248	.304	.273
$P_{t9}$ 極め炊き(炊飯器)	.978	.105	.333	.160
$P_{t10}$ SD1 Merrill(デジカメ)	.922	.137	.238	.174
$P_{t11}$ (CPU)	.891	.295	.532	.380
平均	.942	.153	.263	.180
平均 (all)	.941	.173	.279	.201

かる。ただし、それ以外の製品を対象にしたオープンテストでは、これらのルールの優劣を判定することは難しい。まず、 $P_{t3}$ ,  $P_{t5}$ ,  $P_{t9}$ ,  $P_{t10}$  では、ルール A によって製品対比文と判定された文は一つも無かった(表 4.3 では「-」で示されている)。また、ルール A の F 値が求められた製品について、ルール A, ルール B, ルール A+B を比較すると、優劣関係は製品カテゴリによって変化し、一概にどの手法が有効であるかを論じることはできない。

ルール A では、レビュー文内に製品名が複数出現するかによって製品対比文か否かの判定を行ったが、表 4.3 に示したように製品カテゴリによって精度、F 値に大きな差が見られる。この原因を調べたところ、製品カテゴリによって製品名の記述の仕方に違いがあり、そのために製品名の検出に失敗していることがわかった。競合製品リストの多くが製品名と型番の組(例: AQUOS LC-22K90-W)で表記されており、一方レビュー文内で型番のみが表記されている場合には、本手法では製品名として検出されない。型番のみが記述されている文は、同じメーカーの製品について現行機種と後継機種の対比を行う場合によく見られた。競合製品リストは今回は価格.com の製品リストを基に自動で作成したが、製品カテゴリによって問題があることがわかった。競合製品リストを作成する際、製品と型番の組だけでなく、型番のみもリストに追加することで、この問題はある程度解決できると考えられる。また、同様に製品名として検出が行えなかったケースとして、競合製品がメーカー名や「前使っていた奴」「他の～」といった抽象的な表現で表わされることがある。製品名だけではなく、抽象的な表現もまた対比を行う対象であると判定する必要がある。また、精度に比べ再現率が低いことから、製品名が複数出現した場合には製品対比文である可能性が高まるが、製品対比文の全てを検出する条件としては弱いことを示して

いる。

一方、ルール B では、 $P_{t1}$ =Nexus7 のレビュー文を基に作成した対比表現辞書を用いて製品対比文か否かを判定したが、表 4.4 の「平均」と「平均 (all)」を見ると大きく差が見られなかった。このことから対比表現辞書は様々な製品カテゴリに適用可能なルールとして有効であるといえる。ただし、製品対比文において表 3.1, 3.2 にない対比表現が使われることがあった。例えば、

iPadmini は画面の大きさとアップル社というブランド力が魅力

コレも iPadmini 同様、縦持ちの時の打ちやすさは異常

といった文は製品対比文として検出できなかった。このような検出漏れは、“魅力”、“～やすい”などの対比表現となり得る語句を対比表現辞書に追加することで減らすことができると考えられる。今後、エラー分析などを通じて対比表現辞書を整備・拡充することで製品対比文判定の正解率向上が期待できる。

ルール A やルール B の条件を満たしているレビュー文でも、

Nexus7 買うか iPadmini 買うか

iPad か Android タブレット買うならどっち？

などの文は製品を対比しているわけではないが、誤って製品対比文として判定された。これらの文は対比文と構造が似ているため誤検出された。上記の例は他者に意見を求める文と言えるが、その特徴を分析し、これらの文を検出しないようにルールを洗練したい。

#### 4.2.2 機械学習に基づく手法の評価

機械学習に基づく手法 (3.4 節) では、SVM を学習するために 2 種類の訓練データを利用した。

- 異なる製品カテゴリの文を訓練データとする  
例えば、対象製品が  $P_{t1}$  のとき、 $P_{t2} \sim P_{t11}$  の製品の文を訓練データとする。
- 同じ製品カテゴリの文を訓練データとする  
5 分割交差検定により SVM の学習とテストを行った。

1. の実験結果を表 4.6 に、2. の実験結果を表 4.7 にそれぞれ示す。なお、表 4.7 では 5 分割交差検定における 5 回の試行のマイクロ平均を示した。

4.2.1 項で、製品対比文の特徴は製品カテゴリによって異なると推測した。このことは表 4.6 と表 4.7 で、異なる製品カテゴリのレビュー文を訓練データとした場合と同じ製品カテゴリの文を訓練データとした場合で F 値に大きく差が出ていることから正しいと考えられる。

表 4.6: 実験結果 (SVM; 異なる製品カテゴリの文を訓練データとした場合)

対象製品	A	P	R	F
$P_{t1}^*$ Nexus7(タブレット)	.953	.519	.216	.305
$P_{t2}$ REGZA(液晶テレビ)	.942	.373	.115	.176
$P_{t3}$ うるおい 光クリエール(空気清浄機)	.976	.150	.054	.079
$P_{t4}$ AVR-X4000(AV アンプ)	.978	.102	.250	.145
$P_{t5}$ ブルーレイディーガ(レコーダー)	.984	.050	.042	.045
$P_{t6}$ EH-TW5200(プロジェクタ)	.976	.286	.175	.217
$P_{t7}$ ラムダッシュ(シェーバー)	.956	.242	.066	.103
$P_{t8}$ MDR-DS7500(ヘッドホン)	.954	.313	.267	.288
$P_{t9}$ 極め炊き(炊飯器)	.991	.250	.167	.200
$P_{t10}$ SD1 Merrill(デジカメ)	.948	.131	.090	.107
$P_{t11}$ (CPU)	.935	.471	.281	.352
平均	.964	.237	.151	.171
平均 (all)	.963	.262	.157	.183

表 4.7: 実験結果 (SVM; 同じ製品カテゴリの文を訓練データとした場合)

対象製品	A	P	R	F
$P_{t1}^*$ Nexus7(タブレット)	.955	.549	.394	.458
$P_{t2}$ REGZA(液晶テレビ)	.949	.531	.418	.468
$P_{t3}$ うるおい 光クリエール(空気清浄機)	.983	.594	.339	.432
$P_{t4}$ AVR-X4000(AV アンプ)	.993	.600	.250	.353
$P_{t5}$ ブルーレイディーガ(レコーダー)	.992	.667	.250	.364
$P_{t6}$ EH-TW5200(プロジェクタ)	.980	.451	.404	.426
$P_{t7}$ ラムダッシュ(シェーバー)	.954	.373	.254	.302
$P_{t8}$ MDR-DS7500(ヘッドホン)	.963	.457	.319	.376
$P_{t9}$ 極め炊き(炊飯器)	.993	.333	.167	.222
$P_{t10}$ SD1 Merrill(デジカメ)	.960	.339	.164	.221
$P_{t11}$ (CPU)	.949	.624	.456	.527
平均	.972	.497	.302	.369
平均 (all)	.970	.502	.310	.377

表 4.7 より同じ製品カテゴリのレビュー文を訓練データとしたときの SVM の F 値が 0.369 であるのに対し、表 4.5 のルール A+B の F 値が 0.180 であることから、ルールベースの手法よりも機械学習に基づく手法の方が優れている。これは、機械学習によって製品カテゴリの違いによる対比文の特徴をある程度自動的に学習できるためと考えられる。

全体的に、製品対比文判定の正解率は高いが、精度、再現率、F 値は低い。これは、製品対比文の全体に占める割合が約 3%(表 4.1 より)と低く、製品対比文の正例がデータ内

に多く出現しないことが原因と考えられる。提案手法の開発に用いた製品カテゴリ“タブレット PC(端末)・PDA”ではある程度製品対比文が存在したが、“炊飯器”や“シェーバー”のように製品の寿命が長いために製品の対比があまり行われない製品カテゴリについては対比文の数が少なく、このような製品は特に特徴を学習しづらいことがわかった。そのため、製品対比文の検出の F 値を高めるためには、今回学習素性として用いた製品の出現パターンや対比表現とは異なる素性を用いる必要があると考えている。

次に、個々の学習素性の有効性を調べるために、素性集合から素性を 1 つ除いて SVM を学習し、全素性を用いて学習した SVM と比較した。この実験では、訓練データは同じ製品カテゴリの文とし、5 分割交差検定を行った。結果を表 4.8 に示す。「-」のついている各行は、記号(定義は 3.4 節を参照)で示された素性を除いた素性集合を、最後の行は全素性集合を表す。全素性を用いた場合と比べて、精度 (P), 再現率 (R), F 値 (F) が大きく低下している素性ほど、製品対比文の判定に有効な素性と言える。

表 4.8:  $P_{t1} \sim P_{t11}$  すべてを対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.226	.065	.097
-comExp	.479	.297	.360
-evalExp	.499	.309	.375
-ruleA	.505	.299	.365
-num_P	.499	.306	.373
-num_Y	.509	.308	.376
-evalExp&P	.508	.308	.376
-Y-P_direct	.500	.307	.373
-Y-P_indirect	.501	.303	.371
素性全てを使用	.502	.310	.377

表 4.8 の結果から、最も有効な素性は bow(自立語)であることがわかる。この素性は他の素性に比べて数が多いことから、製品対比文の判定に最も大きく貢献することはある程度予想されたことである。F 値を基準にすれば、bow の次に有効なのは、comExp(対比表現の素性), ruleA(ルール A の素性)であり、それ以外の素性については大差が無かった。これらの素性が bow を除く他の素性と比べて有効であることから、ルールベースの手法で採用されたルールがある程度妥当であるといえる。

同様に、製品カテゴリ毎に素性の有効性を調べた実験の結果を表 4.9~4.19 に示す。これらの表において「-」は製品対比文とされた文が 1 つもないことを表している。bow が一番有効であることは全ての製品で共通しているが、それ以外の素性については製品カテゴリによってその有効性にばらつきがある。例えば、表 4.11 は  $P_{t3}$ (うるおい 光クリエール ACK70N) について 3.3.1 項で提案したルール A の条件を満たすか否かの素性 ruleA が製品対比文の判定に悪影響を与えていることがわかる。また、表 4.17 では、 $P_{t9}$ (極め炊き

NP-BA10) について bow 以外の素性が判定に大きく影響を与えていないことがわかる。このことから、製品対比文をより正確に判定するためには今回提案したものとは異なる素性が必要である。

表 4.9:  $P_{t1}$ =Nexus7 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.529	.232	.322
-comExp	.518	.371	.433
-evalExp	.550	.400	.463
-ruleA	.520	.371	.433
-num_P	.541	.381	.447
-num_Y	.549	.390	.456
-evalExp&P	.554	.390	.458
-Y-P_direct	.532	.394	.453
-Y-P_indirect	.539	.390	.453
素性全てを使用	.549	.394	.458

表 4.10:  $P_{t2}$ =REGZA 26ZP2 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.477	.127	.201
-comExp	.516	.394	.447
-evalExp	.527	.418	.466
-ruleA	.538	.430	.478
-num_P	.530	.424	.471
-num_Y	.550	.430	.483
-evalExp&P	.538	.424	.475
-Y-P_direct	.512	.400	.449
-Y-P_indirect	.504	.406	.450
素性全てを使用	.531	.418	.468

表 4.11:  $P_{t3}$ =うるおい 光クリエール ACK70N を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	-	-	-
-comExp	.545	.321	.404
-evalExp	.571	.357	.440
-ruleA	.613	.339	.437
-num_P	.594	.339	.432
-num_Y	.625	.357	.455
-evalExp&P	.613	.339	.437
-Y-P_direct	.594	.339	.432
-Y-P_indirect	.613	.339	.437
素性全てを使用	.594	.339	.432

表 4.12:  $P_{t4}$ =AVR-X4000 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.400	.083	.097
-comExp	.600	.250	.353
-evalExp	.667	.250	.364
-ruleA	.667	.167	.267
-num_P	.600	.250	.353
-num_Y	.625	.208	.312
-evalExp&P	.667	.250	.364
-Y-P_direct	.667	.250	.364
-Y-P_indirect	.667	.250	.364
素性全てを使用	.600	.250	.353

表 4.13:  $P_{t5}$ =ブルーレイディーガ DMR-BWT650 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	-	-	-
-comExp	.625	.208	.312
-evalExp	.600	.250	.353
-ruleA	.667	.250	.364
-num_P	.667	.250	.364
-num_Y	.667	.250	.364
-evalExp&P	.667	.250	.364
-Y-P_direct	.667	.250	.364
-Y-P_indirect	.667	.250	.364
素性全てを使用	.667	.250	.364

表 4.14:  $P_{t6}$ =EH-TW5200 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	-	-	-
-comExp	.451	.404	.426
-evalExp	.460	.404	.430
-ruleA	.442	.404	.422
-num_P	.451	.404	.426
-num_Y	.462	.421	.440
-evalExp&P	.442	.404	.422
-Y-P_direct	.442	.404	.422
-Y-P_indirect	.440	.386	.411
素性全てを使用	.451	.404	.426

表 4.15:  $P_{17}$ =ラムダッシュ ES-ST25 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.286	.033	.059
-comExp	.369	.254	.301
-evalExp	.338	.221	.267
-ruleA	.354	.230	.279
-num_P	.373	.254	.302
-num_Y	.368	.230	.283
-evalExp&P	.346	.221	.270
-Y-P_direct	.361	.246	.293
-Y-P_indirect	.333	.221	.266
素性全てを使用	.373	.254	.302

表 4.16:  $P_{18}$ =MDR-DS7500 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.200	.015	.028
-comExp	.438	.311	.364
-evalExp	.457	.311	.370
-ruleA	.466	.304	.368
-num_P	.452	.311	.368
-num_Y	.456	.304	.364
-evalExp&P	.456	.304	.364
-Y-P_direct	.452	.311	.368
-Y-P_indirect	.472	.311	.375
素性全てを使用	.457	.319	.376

表 4.17:  $P_{t9}$ =極め炊き NP-BA10 を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	-	-	-
-comExp	.333	.167	.222
-evalExp	.333	.167	.222
-ruleA	.333	.167	.222
-num_P	.333	.167	.222
-num_Y	.333	.167	.222
-evalExp&P	.333	.167	.222
-Y-P_direct	.333	.167	.222
-Y-P_indirect	.333	.167	.222
素性全てを使用	.333	.167	.222

表 4.18:  $P_{t10}$ =SD1 Merrill を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	-	-	-
-comExp	.295	.148	.197
-evalExp	.339	.156	.213
-ruleA	.339	.156	.213
-num_P	.339	.164	.221
-num_Y	.338	.180	.235
-evalExp&P	.339	.156	.213
-Y-P_direct	.322	.156	.210
-Y-P_indirect	.317	.156	.209
素性全てを使用	.339	.164	.221

表 4.19:  $P_{t11}$ =CPU を対象とした実験結果 (SVM; 学習素性ごと)

	P	R	F
-bow	.594	.222	.323
-comExp	.580	.444	.503
-evalExp	.642	.462	.537
-ruleA	.615	.468	.532
-num_P	.605	.421	.497
-num_Y	.624	.456	.527
-evalExp&P	.636	.480	.547
-Y-P_direct	.622	.462	.530
-Y-P_indirect	.627	.462	.532
素性全てを使用	.624	.456	.527

### 4.3 製品対比文における製品名の出現の有無に関する考察

製品対比文の判定を行うとき、レビュー文内に製品名が出現する場合としない場合では、後者の方が文脈や製品知識などの情報が必要であり、より判定が難しいと考えられる。そこで  $P_{t1}$ =Nexus7のレビュー文に、製品対比文であるか否かのタグ付けに加え、“TARGET”及び“PRODUCT”というタグを追加した。それぞれのタグの定義を以下に示す。

- TARGET  
対象製品名が文中に出現するか否かを示すタグである
  - TARGET=2 のとき、製品対比文内に対象製品名が現れる
  - TARGET=1 のとき、製品対比文内に対象製品名は現れないが、文脈上には現れる
- PRODUCT  
競合製品名が文中に出現するか否かを示すタグである
  - PRODUCT=2 のとき、製品対比文内に競合製品名が現れる
  - PRODUCT=1 のとき、製品対比文内に競合製品名は現れないが、文脈上には現れる
  - PRODUCT=0 のとき、製品対比文内にも文脈上にも競合製品名が現れない

PRODUCT=0である対比文の例を以下に示す。対象製品はNexus7である。

- “Appleの商品と比べると見劣りするけど十分合格点”
- “以前所有していた端末は文字入力からして全体的にモッサリしていましたが、Nexus7は問題なくキビキビ動いており大満足です”

そして、全ての製品対比文を正例としたときと、製品名が文中に出現する文、すなわちTARGET=2またはPRODUCT=2である文のみを正例としたときで、実験の結果を比較した。今回はルールベースの手法の正解率、精度、再現率、F値を調べた。

実験結果を表4.20、表4.21に示す。

製品対比文を抽出する問題の難しさとして、(I)全ての対比文を検出する場合(表4.20)と(II)製品名が出現する対比文のみを検出する場合(表4.21)を比べると(I)の方が困難であると予想した。ルールA+Bの結果を比較すると、再現率では(I)が0.504、(II)で0.640と予想通りの結果が得られた。これに対し精度では(I)が0.416と(II)の0.241を上回った。これはルールBが製品名ではなく評価表現を基に製品対比文を判定しており、(II)では誤検出となる製品名を含まない対比文を取得できているためである。

表 4.20: 製品対比文全てに対しての評価

	A	P	R	F
ルール $A$	.955	.623	.192	.294
ルール $B$	.941	.408	.424	.416
ルール $A + B$	.941	.416	.504	.456

表 4.21: 製品名が文内に出現するもののみを製品対比文とした場合の評価

	A	P	R	F
ルール $A$	.981	.597	.404	.482
ルール $B$	.948	.208	.474	.289
ルール $A + B$	.947	.241	.640	.350

## 第5章 結論

### 5.1 まとめ

本論文では、複数の製品の対比を行う評判情報に着目し、製品対比文を自動的に判定する手法を提案し、その有効性を実験的に検証した。

まず Nexus7 に関するレビュー文を分析、精査することで、製品対比文における製品名の出現パターンと、対比文を行う文章に特徴的に現れる表現が存在することを発見し、製品対比文が満たすべき条件をチェックするルール設計を行った。作成したルールを用いて製品対比文の判定を行うシステムを実装した。しかし、ルール設計を行った Nexus7 以外の製品カテゴリに対して同様に判定を行ったところ、判定の F 値は低下し、製品対比文に出現する特徴は製品カテゴリ毎に異なることも明らかになった。また、レビュー文内に現れる製品名の記述の仕方にはルール A で想定していないものもあることがわかった。

次に、製品対比文の判定を行うのに有効であると考えられる素性を考案し、機械学習によって製品対比文か否かの判定を行うモデルを学習した。同じ製品カテゴリのレビュー文を訓練データとした場合(表 4.7)に比べ、異なる製品カテゴリのレビュー文を訓練データとした場合(表 4.6)の評価指数は大きく劣った。このことから、製品カテゴリによって製品対比文の持つ特徴が異なることがわかった。

個々の素性の有効性を調べたところ、文中の自立語、対比表現の有無、ルール A(複数の製品を含むか)については有効性は確認されたものの、他の素性は製品対比文判定の F 値の向上に大きく貢献しなかった。また、製品カテゴリによって素性の有効性にばらつきが見られた。従って、対比文の判定に有効で、かつどのような製品カテゴリに対しても有効に働く素性を新たに探求する必要がある。

### 5.2 課題

今後は、製品対比文の判定の精度・再現率の改善を目指す。レビュー文に現れる製品名の検出に用いる競合製品リストの作成の際、ウェブサイトの製品名リストから単純に取得するだけでなく、型番や略称などの他の呼称も考慮する。また、対比表現の精査、拡充を行うことで判定の精度改善を図るとともに、製品カテゴリに依らない製品対比文の判定を行う手法についても検討したい。

また、本論文では製品対比文の判定を文単位で行った。しかし、“iPad2 は重量が 613g. iPad mini は 312g.” といったように、文単位ではそれぞれの製品についての特徴を述べ

ているだけだが，2つの文では対比を行っている場合が多く存在した．このような文も本来製品対比文として検出を行うべきである．文脈を考慮した製品対比文の判定手法についても検討をしたい．

最後に，検出した製品対比文をユーザにどのように分かりやすく提示するかも重要な課題である．複数の製品対比文を見やすく並べたり，要約する手法を検討する必要がある．

## 参考文献

- [1] 乾孝司, 奥村学. テキストを対象とした評価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.
- [2] 立石健二, 石黒義英, 福島俊一. インターネットからの評判情報検索. 情報処理学会 研究報告, NL-144-11, pp75-82, 2001.
- [3] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. 意見抽出のための評価表現の収集. 自然言語処理, Vol.12, No.2, pp.203-222, 2005.
- [4] Xiaowen Ding, Bing Liu, Philip S. Yu. A Holistic Lexicon-Based Approach to Opinion Mining. WSDM'08 Proceedings of the 2008 International Conference on Web Search and Data Mining, pages231-240, 2008. Pages 231-240
- [5] Vasileios Hatzivassiloglou, Kathleen R. McKeown. Predicting the semantic orientation of adjectives. ACL '98 Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics, Pages 174-181 , 1997.
- [6] 倉島健, 別所克人, 戸田浩之. 比較評価情報に基づくランキング手法. 日本データベース学会 Letters Vol.6, No.1, 2007.
- [7] Murty Ganapathibhotla and Bing Liu. Mining Opinions in Comparative Sentences. Proceedings of the 22nd International Conference on Computational Linguistics, pages241-248, 2008.
- [8] Nitin Jindal and Bing Liu. Mining comparative sentences and relations. Proceedings of the AAAI, pp.1331-1336, 2006.
- [9] Kaiquan Xu, Stephen Shaoyi Liao, Jiexun Li, and Yuxia Song. Mining comparative opinions from customer reviews for competitive intelligence. Decision Support Systems, Vol. 50, No. 4, pp. 743-754, 2011.
- [10] 工藤 拓, 松本 裕治. チャンキングの段階適用による日本語係り受け解析. 情報処理学会論文誌 Vol.43, No.6, pp.1834-1842, 2002.