

商品レビューの網羅性と支持度の関係の分析

別所 暁秀¹ 湯本 高行¹ 新居 学¹ 佐藤 邦弘¹

概要：

商品を購入する際、通販サイトの商品レビューを閲覧すれば有用な情報を獲得できる。しかし、レビューに記述された意見は必ずしも有用ではない。本研究では、有用なレビューと有用でないレビューを予測するために、支持度と意見の網羅性に関する分析を行う。本研究では、意見の網羅性を表すための指標として網羅度と重み付き網羅度を定義する。網羅度は、1件のレビューに記述された意見がどの程度全体の意見を網羅しているかを表す。また、重み付き網羅度は、意見中の語の出限頻度で網羅度を重み付けした指標である。支持度とはレビューに対する閲覧者の投票によって算出される、支持の程度を表す指標である。本研究では、対象レビューを支持度と網羅度および重み付き網羅度の値によって分類する事で、支持度と網羅性の関係を明らかにする。

1. はじめに

近年、インターネットが普及したことにより、商品を購入する際に Web を利用する機会が増加している。Web を利用すれば商品の評判情報を容易に獲得できるため、商品購入の意思決定に役立つ。Web から商品に関する評判情報を獲得するための手段として、EC サイト等で提供されている商品レビュー（レビュー）の閲覧があげられる。レビューには投稿者による商品に関する記述（意見）が掲載されており、複数の意見を閲覧すれば評判情報を獲得できる。

これらの背景から、ある商品に関する評判情報を抽出するために意見を要約する研究が盛んに行われている [1][2][3]。それらの先行研究では、商品に対して投稿された全てのレビューから商品の機能や特徴を表す語を抽出し、その語に対する評判情報を要約している。しかし、全てのレビューから評判情報を抽出しているため、閲覧者から支持されていない情報も含んでしまう。情報の有用性を考慮した場合、支持されていない情報は要約をする際に不適切であると考えられる。

このような問題点があることから、意見の有用性に注目した研究が行われている。意見の有用性によりレビューを分類する研究として、Liu ら [4] の研究がある。Liu らは、複数の指標を用いて、有用なレビューを分類する手法を提案している。まず、学習データとして多数のレビューを用意し、意見の有用性を人手により 1 件ずつ 4 段階で評

価する。さらに、4 段階で最も評価の低いレビューを低評価レビュー、その他のレビューを高評価レビューとし、複数の指標を用いてそれら 2 つの評価のレビューを分類している。

本研究では、意見の網羅性に注目することで、支持を集めるレビューと支持されていないレビューを分類することを目的とする。ここで、分類を行うための指標（分類指標）として、意見の網羅性の度合いを表す網羅度と、意見中の語の出限頻度で網羅度を重み付けした重み付き網羅度という 2 つの指標を定義した。さらに、レビューの有用性を数値化するために、レビューに対する閲覧者の投票を用いた。

レビューの分類実験では、支持度と分類指標の値によってレビューを分類できるか実験した。また、提案手法の比較対象となる分類指標を定義し、提案手法の優位性を検証した。比較対象の分類指標として、意見の文書長から算出される指標と、簡潔さの度合いを表す指標である簡潔度を定義し、分類の精度を比較した。

2. 各指標の算出

本章では、本研究で使用するレビューの構成要素について説明し、その要素から算出される各指標について説明する。

2.1 レビューの構成要素

本研究では、Amazon.co.jp^{*1}で提供されているレビューを支持を集めるレビューと支持されていないレビューに分

¹ 兵庫県立大学大学院工学研究科

^{*1} <http://www.amazon.co.jp/>

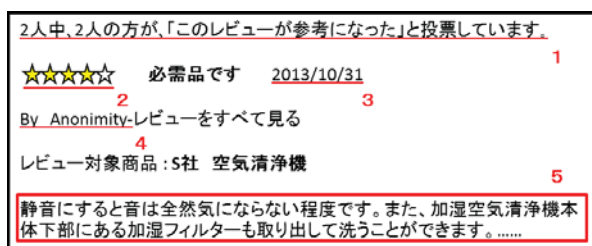


図 1 Amazon.co.jp のレビュー

類する。レビューの概要図を図 1 に示す。

ここで、図 1 のような商品に関する 1 件の記事をレビューと定義する。また、レビューは図 1 に示すように以下の要素で構成されているとする。

1. 閲覧者によるレビューに対する投票
2. 商品に対する 5 段階評価
3. 投稿日時
4. 投稿者名
5. 意見

上記の番号は図中の赤字に対応している。これら 5 つの要素のうち、閲覧者によるレビューに対する投票により、そのレビューがどの程度閲覧者から支持されているか算出する。また、意見を対象としたテキストマイニングを行うことで分類指標を算出する。

本稿では、 N 件のレビューが投稿された商品があった場合、各レビューを r_i 、 r_i の意見を a_i と定義する ($i = 1, 2, \dots, N$)。また、全ての意見の集合 (意見集合) を A と表記する。

2.2 レビューの支持度

図 1 に示すように、Amazon.co.jp ではレビューの閲覧者がその記事の内容が参考になったか否か投票を行う事ができる。ここで、1 件のレビューに対する投票について考える。まず、多数の閲覧者がレビューに対して「参考になった」と投票している場合、そのレビューは多くの支持を受けていると言える。一方、少数の閲覧者が「参考になった」と投票している場合も支持はされているが、前者よりも支持の度合いは低いと考えられる。以上より、閲覧者の投票における投票総数と「参考になった」と投票した数を考慮することで支持の度合いを算出することができる。ここで、レビューに対する支持の度合いを支持度と定義する。あるレビュー r_i の支持度 $sup(r_i)$ は式 1 により算出する。

$$sup(r_i) = vote_{good}(r_i) - (vote_{all}(r_i) - vote_{good}(r_i)) \quad (1)$$

式 1 において、 $vote_{all}(r_i)$ は r_i に対して投票を行った人数であり、 $vote_{good}(r_i)$ は r_i が参考になったと投票した人数である。ここで、 $sup(r_i) > 0$ のレビューを支持レビュー、 $sup(r_i) < 0$ のレビューを不支持レビューとする。また、投票が行われていないレビューは $sup(r_i) = 0$ とする。

2.3 レビューの網羅度

レビューの構成要素の 1 つである意見には、商品に関する記述がされている。本研究では、分類指標を算出するために文書の網羅性に着目している。ここで、文書の網羅性とは、ある文書の集合があったとき、文書 1 件が集合全体で出現する話題と同様の話題をどの程度含んでいるか表す。

例として、掃除機に関するレビューの集合 (レビュー集合) の要素、レビュー X 、 Y を例にして網羅性について説明する。レビュー X 、 Y それぞれの意見を以下に示す。

X : 吸引力が良いです。また、音も静かで使い勝手も良いです。

Y : 購入前から噂で聞いていましたが、吸引力が凄いです。

X 、 Y それぞれの文字数はほぼ等しいが、 X では「吸引力」、「音」、「使い勝手」に関する話題が記述されており、 Y では「吸引力」に関する話題のみ記述されている。よって、レビュー X は Y よりも網羅性が高い。また、網羅性の高い意見ほど、商品に関する情報量が多いことがわかる。

以上の例より、本研究では支持度の高いレビューの意見ほど得られる情報が多い、つまり網羅性が高いほど支持度が高いと考える。また、網羅性の度合いを算出するために意見中に出現する特徴語に着目する。

2.3.1 特徴語の定義

特徴語とは、意見に出現する商品に関する機能や特徴、または商品自体を表す語である。本研究では、意見に出現する語のうち、以下の品詞情報を持つ語を特徴語と定義する。

- 「名詞・一般」
- 「名詞・ナイ形容詞語幹」

この品詞情報は形態素解析器 MeCab*2 を用いて得られる結果である。さらに、以下に示す条件に当てはまる語の出現パターンが確認できた場合、それらを連結処理する事で複合名詞として取り扱い、特徴語として抽出する。

- 「名詞・非自立」を除く名詞の連続
- 「接頭詞・名詞接続」の語の次に名詞が出現

2.3.2 網羅度の算出

前述のように、本研究では支持度の高いレビューほど全体の意見を網羅した内容を記述していると考えられる。そこで、レビューがどの程度全体の意見を網羅しているかを表す指標を網羅度と定義する。レビュー集合における全意見集合 A 中に出現する特徴語の集合を W_A 、レビュー集合の要素である 1 件のレビューの意見 a_i に出現する特徴語の集合を W_i とし、 a_i の網羅度 $coverage(r_i)$ を式 2 によって定義する。

*2 <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

$$coverage(r_i) = \frac{|W_i|}{|W_A|} = \frac{\sum_{w \in W_i} 1}{\sum_{w \in W_A} 1} \quad (2)$$

式2における w は特徴語である。また、 W_A と W_i の関係性は式3で表される。

$$W_A = \bigcup_{i=1}^N W_i \quad (3)$$

N は各商品のレビュー件数であり、 W_A は各商品ごとに固有である。

本研究では、多くの意見で出現する特徴語ほど商品に関する重要な語であると考えられる。そこで、特徴語の出現頻度の値により網羅度を重み付けする。特徴語の出現頻度はDF値によって算出する。ここで、特徴語 w のDF値は、意見の集合 A において w が出現する意見の件数を表し、 $DF(w, A)$ と表記する。式2をDF値により重み付けした網羅度 $weightCov(r_i)$ を次式によって定義する。

$$weightCov(r_i) = \frac{\sum_{w \in V_i} DF(w, A \setminus \{a_i\})}{\sum_{w \in V_A} DF(w, A \setminus \{a_i\})} \quad (4)$$

ここで、式4の $A \setminus \{a_i\}$ は a_i を除く意見の集合である。

以降、 $coverage(r_i)$ を網羅度、 $weightCov(r_i)$ を重み付き網羅度とよぶ。

3. 実験

本章では、重み付き網羅度によって支持レビューと不支持レビューを分類できるか実験する。まず、網羅度と重み付き網羅度による分類をそれぞれ行い、分類結果を比較することで特徴語に重み付けを行うことが有効であるか検証する。次に、重み付き網羅度のベースラインとして分類指標を複数用意し、それらの分類結果と精度を比較することで重み付き網羅度の優位性を検証する。

3.1 実験データ

本実験では、Amazon.co.jpで販売されている21商品(対象商品)を実験データとした。ここで、対象商品を選択する際にレビュー数が50件以上であることを条件とした。これは、レビュー数の多い商品では閲覧者による投票が行われたレビューが多く投稿されていると考えたためである。

さらに、各商品のレビュー集合において、支持度の上位/下位それぞれ10%のレビューを分類対象のレビュー(以下、対象レビュー)として取得した。ここで、対象レビュー以外のレビューは、投票数が少ないものとレビューに対する評価が分かれているものである。そのようなレビューは、今後の投票により支持度の正負が変わる可能性があり、実験データとしてふさわしくないため対象レビューから除外

した。各対象商品を表すカテゴリ名を表1に示し、各対象商品に投稿されたレビュー数と対象レビュー数を表2に示す。表2のレビュー数は2013年6月13日時点のデータである。以降、商品を表現する際にカテゴリ名を使用する。

表1 対象商品のカテゴリ名

商品名	カテゴリ名
dyson DC26THCOM	掃除機
任天堂 マリオパーティ 8	TV ゲーム
CASIO MTP-1239DJ-2AJF	腕時計
Bauhutte BM-39 Arancia	オフィスチェア
Canon EOS Kiss X5	デジカメ
SHARP KC-A50-W	空気清浄機
SONY NW-S754	ウォークマン
YUPITERU YPL502si	カーナビ
EPSON EP-802A	インクジェット複合機
T-fal BF802022A	電気ケトル
小川糸著 食堂かたつむり	小説
SAVAS アクア ホエイプロテイン 100	サプリメント
Creative EP-630i	イヤホン
PHILIPS PT725	シェーバー
ぺんてる PG1005	シャープペン
バイオ お風呂のカビきれい	洗剤
LOGICOOL M570	マウス
MEYER MPC-2.3RD	圧力鍋
クリエイティブ SP-SBS-A120	スピーカー
ELECOM DE-U01L-4710	モバイルバッテリー
SHARP PW-M800	電子辞書

表2 対象商品とそのレビュー数

カテゴリ名	レビュー総数	対象レビュー数
掃除機	82	16
TV ゲーム	89	18
腕時計	136	27
オフィスチェア	109	22
デジカメ	127	25
空気清浄機	306	61
ウォークマン	121	24
カーナビ	288	58
インクジェット複合機	160	32
電気ケトル	142	28
小説	216	43
サプリメント	227	45
イヤホン	122	24
シェーバー	212	42
シャープペン	64	13
洗剤	128	26
マウス	510	102
圧力鍋	115	23
スピーカー	372	74
モバイルバッテリー	87	17
電子辞書	178	36

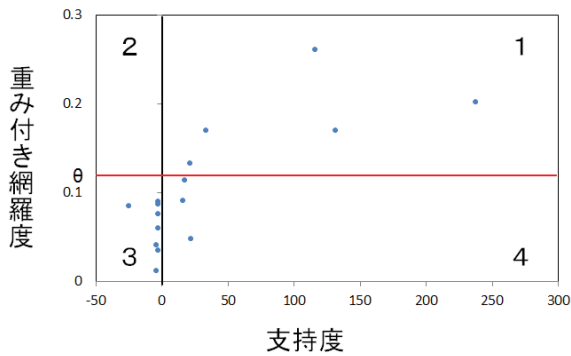


図 2 支持度と分類指標に関する散布図

3.2 レビューの分類実験

各商品ごとに以下の手順でレビューの分類実験を行った。商品のレビュー集合について 2.2 節の手法により全レビューの支持度を算出し、支持度上位/下位 10% ずつのレビューを対象レビューとして抽出した。次に 2.3 節の手法により各対象レビューごとに重み付き網羅度を算出した。

次に、図 2 に示すように散布図を使用することで分類の精度検証を行った。ここで、図 2 を例にして分類の手法を述べる。

まず、支持度と重み付き網羅度それぞれでしきい値を設ける。さらに、支持度のしきい値から縦軸に平行な補助線と、重み付き網羅度のしきい値から横軸に平行な補助線を引くことで、散布図を図 2 に示すように 4 つの象限に分類する。ここで、支持度のしきい値は 0、重み付き網羅度のしきい値は θ である。よって各象限の支持度および重み付き網羅度の値は表 3 に示すようになる。

象限	支持度	重み付き網羅度
1	> 0	$> \theta$
2	> 0	$\leq \theta$
3	≤ 0	$\leq \theta$
4	≤ 0	$> \theta$

表 3 において、しきい値 θ は全レビューの重み付き網羅度の算術平均である。また、3.2.1 節以降で説明する、比較対象の分類指標の場合のしきい値も各指標の算術平均とした。本研究では支持度が高いレビューほど分類指標の値が高くなると考えているため、表 3 および図 2 より支持レビューが第 1 象限に、不支持レビューが第 3 象限に分布していれば正しく分類できたとと言える。ここで、分類の精度を算出するための尺度として正答率を用いた。正答率は全データのうち分類を正しく行うことができたデータの割合を表す尺度であり、正答率が 1 に近いほど分類の性能が良いと言える。正答率 *accuracy* は式 5 によって算出できる。

$$\frac{\text{第 1 象限のレビュー数} + \text{第 3 象限のレビュー数}}{\text{対象レビュー数}} \quad (5)$$

表 4 重み付き網羅度と網羅度の正答率

カテゴリ名	重み付き網羅度	網羅度
掃除機	0.88	0.75
TV ゲーム	0.69	0.63
腕時計	0.85	0.73
オフィスチェア	0.75	0.75
デジカメ	0.83	0.79
空気清浄機	0.63	0.59
ウォークマン	0.79	0.71
カーナビ	0.66	0.67
インクジェット複合機	0.79	0.69
電気ケトル	0.64	0.68
小説	0.60	0.66
サプリメント	0.72	0.59
イヤホン	0.59	0.63
シェーバー	0.70	0.64
シャープペン	0.83	0.75
洗剤	0.63	0.71
マウス	0.67	0.63
圧力鍋	0.77	0.91
スピーカー	0.59	0.68
モバイルバッテリー	0.56	0.68
電子辞書	0.71	0.53
平均	0.71	0.69

3.2.1 重み付けの有効性検証

以上の手順で分類を行い、網羅度、重み付き網羅度それぞれの指標による正答率を算出することで、重み付けの有効性を検証した。ここで、各分類指標の正答率を表 4 に示す。表 4 において下線が引かれた値は、各カテゴリの正答率の最大値である。表 4 より、21 商品中、12 商品で重み付き網羅度による分類精度が上回り、1 商品で精度が等しくなった。また、正答率の平均値は重み付き網羅度が 0.01 上回った。この結果から、2 指標の精度に差はほぼ見られなかった。しかし、重み付き網羅度の場合、特徴語の集合に存在する不適切な語の影響を減少することができるため、その点では有用である。そのような語の例として誤字が挙げられるが、重み付き網羅度の場合、誤字の DF 値は小さくなることが明らかなため式 4 の算出の際に影響が小さくなる。

3.2.2 分類精度の比較実験

重み付き網羅度の優位性を検証するために、指標の比較対象として意見の文書長と簡潔さの 2 つを分類指標とする実験を行った。ここで、本研究における文書長とは、句読点や記号などを含む文字数を表し、レビュー r_i の文書長 $length(r_i)$ は意見 a_i の文書長と定義する。

また、意見の簡潔さとはある特徴語に対する記述がどれだけ端的に表されているかを示し、より多くの種類の特徴語が含まれ、文書長がより短い意見ほど簡潔であるとする。ここで、簡潔さの度合いを簡潔度とし、簡潔度

表 5 重み付き網羅度, 文書長, 簡潔度それぞれの正答率

カテゴリ名	重み付き網羅度	文書長	簡潔度
掃除機	<u>0.88</u>	0.81	0.44
TV ゲーム	<u>0.69</u>	0.63	0.50
腕時計	<u>0.85</u>	0.77	0.58
オフィスチェア	<u>0.75</u>	<u>0.75</u>	0.35
デジカメ	<u>0.83</u>	<u>0.83</u>	0.46
空気清浄機	<u>0.63</u>	0.61	0.48
ウォークマン	<u>0.79</u>	0.71	0.58
カーナビ	0.66	<u>0.69</u>	0.57
インクジェット複合機	<u>0.79</u>	0.71	0.48
電気ケトル	<u>0.64</u>	0.57	0.50
小説	<u>0.60</u>	0.55	0.45
サプリメント	<u>0.72</u>	0.65	0.59
イヤホン	<u>0.59</u>	<u>0.59</u>	0.45
シェーバー	<u>0.70</u>	0.63	0.58
シャープペン	<u>0.83</u>	0.67	0.58
洗剤	0.63	0.58	<u>0.75</u>
マウス	<u>0.67</u>	0.63	0.48
圧力鍋	0.77	<u>0.91</u>	0.50
スピーカー	0.59	<u>0.67</u>	0.49
モバイルバッテリー	0.56	<u>0.69</u>	0.44
電子辞書	<u>0.71</u>	0.50	0.62
平均	<u>0.71</u>	0.67	0.52

$simplicity(r_i)$ を式 6 によって定義する.

$$simplicity(r_i) = \frac{|W_i|}{length(r_i)} \quad (6)$$

ここで, 式 6 の $|W_i|$ は意見 a_i に出現する特徴語の種類数である.

以上 2 つの分類指標と重み付き網羅度の分類精度を比較する. 各分類指標の正答率を表 5 に示す. 表 5 において下線が引かれた値は, 各カテゴリの正答率の最大値である. 表 5 より, 重み付き網羅度と簡潔度を比較した場合, 21 商品中, 20 商品で分類精度が上回った. また, 重み付き網羅度が正答率の平均で 0.19 上回った. よって, 簡潔度と比較した場合, 分類精度が高いことがわかった.

次に, 文書長と比較した場合, 21 商品中, 14 商品で分類精度が上回り, 3 商品で精度が等しくなった. また, 正答率の平均を比較した場合, 網羅度による分類が 0.04 上回っていた. よって, 重み付き網羅度による分類の場合が精度が良いことはわかったが, 顕著な差は見られなかった.

重み付き網羅度の正答率を見た場合, 小説で数値が低い結果となった. これは, 意見中で話題となるのは特定のシーンに関する感想などであり, 特徴語の表記揺れが多く発生し網羅度の算出に影響したと考えられる.

また, 比較実験において, 簡潔度と比較した場合には優位性が検証できたが, 文書長と比較した場合に顕著な差が見られなかった. そこで, 2 つの指標の分類の差異を知るために, 重み付き網羅度による分類に成功し, 文書長によ

る分類に失敗したレビューの特徴分析を行った. まず, 対象レビューは 2 つの分類指標それぞれで成功したか否かによって 4 つのグループに割り振ることができる. それぞれのグループにおけるレビュー数を表 6 に示す.

表 6 各グループのレビュー数

		重み付き網羅度	
		成功	失敗
文書長	成功	430	67
	失敗	90	169

表 6 より, 重み付き網羅度による分類に成功し, 文書長による分類に失敗したレビュー数は 90 件であった. また, 90 件中, 支持レビューは 49 件であり, 不支持レビューは 41 件であった.

この 90 件のレビューの特徴を分析するために, 人手により意見を閲覧したところ, 表 7 に示す 4 パターンの記述内容が多いことがわかった.

表 7 意見の記述内容の分類

分類	特徴
不体裁	意見中に顔文字やネットスラングを含む. 誤字が多い.
無関係	文書の大半で対象商品に対する意見が記述されていない.
不具合	対象商品の不具合に関する記述がされている.
簡潔	商品の機能や特徴について簡潔に記述されている.

支持レビュー, 不支持レビューそれぞれで表 7 の特徴にあてはまるレビュー数を表 8 に示す.

表 8 重み付き網羅度による分類に成功し, 文書長による分類に失敗したレビューの記述内容の内訳

	不体裁 無関係 不具合 簡潔 その他					総数
	不体裁	無関係	不具合	簡潔	その他	
支持	2	0	6	31	10	49
不支持	15	10	7	2	7	41

表 8 より不支持レビューにおいて不体裁と無関係が合わせて 25 件 (61%) あったため, 不支持レビューは体裁の悪いレビューが多いことがわかった. これら 2 パターンの記述はともに情報を獲得する際に無駄な情報を多く含むという共通点がある.

また, 支持レビューでは簡潔に分類されたレビューが 64% であった. そこで, 分類指標である意見の簡潔度に注目することで支持レビューの特徴を分析した. 各商品ごとに全レビューの簡潔度を算出し, 簡潔度の算術平均の値と比較した. 支持レビュー, 不支持レビューそれぞれにおける簡潔度の平均値以上のレビュー数とその割合を表 9 に示す.

表 9 より, 支持レビューにおいて簡潔度の平均値以上のレビューが 73% であることが分かった.

以上の分析より, 重み付き網羅度による分類に成功し, 文書長による分類に失敗したレビューにおいて, 支持レ

表 9 重み付き網羅度による分類に成功し、文書長による分類に失敗したレビューの簡潔度との関係

	レビュー総数	平均以上のレビュー数	割合
支持	49	36	0.73
不支持	41	18	0.44

ビューは簡潔度の高いレビューが多く、不支持レビューは無駄な情報を多く含む意見が多いことが分かった。よって、重み付き網羅度による分類は、支持レビューのうち、網羅性の高い意見だけではなく、簡潔に記述されたものを発見することに優れていることがわかった。

次に、重み付き網羅度による分類に失敗し、文書長による分類に成功したレビューの特徴を分析することで重み付き網羅度の改善すべき点を調査した。この条件に当てはまるレビューは、表 6 より 67 件であった。また、67 件中、支持レビューは 27 件であり、不支持レビューは 40 件であった。このレビューについて、上述と同様の分析を行った。まず、記述内容に関する分析を行った結果を表 10 に示す。

表 10 重み付き網羅度による分類に失敗し、文書長による分類に成功したレビューの記述内容の内訳

	不体裁	無関係	不具合	簡潔	その他	総数
支持	4	2	9	5	7	27
不支持	8	13	5	7	7	40

表 10 より、不支持レビューにおいて不体裁と無関係が合わせて 21 件 (53%) であった。さらに、21 件中の多くのレビューで価格に関する記述がされていた。価格は商品を購入する際の重要な要素である。また、意見中に記述されることが多いため DF 値が高くなる。しかし、価格をレビューを閲覧せずとも容易に知ることができるため、レビューから獲得できる情報としては重要ではない。つまり、DF 値が高いが重要ではない語を含むと分類精度が低下することがわかった。

次に、簡潔度の平均との比較結果を表 11 に示す。

表 11 重み付き網羅度による分類に失敗し、文書長による分類に成功したレビューの簡潔度との関係

	レビュー総数	平均以上のレビュー数	割合
支持	27	14	0.50
不支持	40	23	0.57

表 11 より、支持レビュー、不支持レビューどちらの場合も簡潔度の平均以上のレビューの割合は 5 割程度となり、特徴は見られなかった。

以上の分析より、重み付き網羅度による分類は、DF 値が高いが重要ではない語を含むレビューがある場合に精度が低下することがわかった。今後、そのような語を特徴語から除外することで分類精度の向上が期待できる。

4. おわりに

本稿では、支持レビューと不支持レビューを分類するために、分類指標として網羅度と重み付き網羅度を定義し分類を行った。また、レビューがどの程度支持されているか表す指標として支持度を定義した。実験では、提案した 2 つの分類指標の分類結果を比較し、重み付き網羅度がより有効であることを示した。また、重み付き網羅度の優位性を検証するため、比較対象として文書長と簡潔度による分類を行った。重み付き網羅度による分類を行った結果、分類精度は 0.71 であった。また、文書長と簡潔度による精度はそれぞれ 0.67 と 0.52 であった。この結果から、簡潔度に対する優位性は示せたが、文書長との差はほぼ見られなかった。そこで、分類結果を利用して重み付き網羅度による分類の特徴を分析した。結果として、支持レビューのうち、網羅性の高い意見だけではなく、簡潔に記述されたものを発見することに優れていること、また、DF 値が高いが重要ではない語を含むと分類精度が低下することがわかった。今後は、DF 値が高いが重要ではない語を特徴語から除外する手法について検討する。

謝辞 本研究の一部は、平成 25 年度科研費若手研究 (B) 「情報の詳細関係に基づく Web ページの組織化」(課題番号: 24700097) によるものである。

参考文献

- [1] 平山拓央, 湯本高行, 新居学, 佐藤邦弘: 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, 第 155 回情報処理学会研究報告, データベース・システム研究会報告, 2012-DBS-155(3), pp.1-9 (2012).
- [2] Mingqing Hu, Bing Liu: Mining and Summarizing Customer Reviews, Knowledge Discovery and Data Mining(KDD)-2004, pp.22-25 (2004).
- [3] 谷本融紀, 太田学: 評価表現辞書を用いた評判情報の極性値計算, DEIM Forum2012, D2-3 (2012).
- [4] Jingjing Liu, Yunbo Cao, Chin-Yew Lin, Yalou Huang, Ming Zhou: Low-Quality Product Review Detection in Opinion Summarization, Empirical Methods in Natural Language Processing(EMNLP), pp. 334-342 (2007).