

統計的手法を用いたレビューポートフォリオの構築

小西 卓哉[†] 手塚 太郎^{††} 木村 文則^{†††} 前田 亮^{††}

^{†,††,†††} 立命館大学情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: [†] cm005069@ed.ritsumei.ac.jp, ^{†††} fkimura@is.ritsumei.ac.jp

^{††} {tezuka,amaeda}@media.ritsumei.ac.jp

あらまし Web 上の商品レビューサイトを利用する際、特定の意見のみを参照すると偏った情報により誤った判断をするリスクがある。そこで各商品に対し多面的な角度からレビューを提示することにより、誤った判断を行うリスクを軽減する手法としてレビューポートフォリオの構築を提案する。統計的手法に基づいたレビューの組み合わせを提示することで、様々な意見を反映した信頼性の高い情報を提供する枠組みを示す。とくに本稿では階層的クラスタリングによるレビューの分類を行い、周辺情報から生成したルールを用いて行うレビューの選択の手法について記す。

キーワード レビュー, 情報の信頼性, 情報抽出

“Review Portfolio” by Statistical Methods

Takuya KONISHI[†] Taro TEZUKA^{††} Fuminori KIMURA^{†††} Akira MAEDA^{††}

^{†,††,†††} College of Information Science and Engineering, Ritsumeikan University

1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan

E-mail: [†] cm005069@ed.ritsumei.ac.jp, ^{†††} fkimura@is.ritsumei.ac.jp

^{††} {tezuka,amaeda}@media.ritsumei.ac.jp

Abstract When one uses product review sites on the Internet, there is a risk that he/she makes an inappropriate decision by considering only to a particular opinion. We therefore propose a method of building “review portfolio” in order to reduce such risk by presenting reviews consisting of different view point. We also introduce a framework that provides reliable information considering a variety of reviews based on statistical methods. Specially, we create review portfolio by classifying reviews using scores given to different aspects of the item and selecting a review from each cluster based on several rules for extracting useful reviews.

Keyword Review, Reliability of Information, Information Extraction, Classification Model

1. はじめに

近年 Web を介して個人からの情報発信が普及したことで、個人が商品やサービスに対して評価を行うレビュー情報の入手が容易になった。これらレビューは一般ユーザの商品購買に大きな影響を及ぼすため、Amazon.co.jp[1]や価格.com[2]のような商品販売支援サイトはユーザからレビューの提供を募り公開している。また企業の商品開発においてもレビュー情報などのユーザの意見や動向を利用する手法が注目を集めている。

レビューにおけるように個人レベルの意見や評価を入手することが容易になった一方、個人から発信される情報の問題点も指摘されている。そのひとつとして、個人から一方的に発信される情報であるため、信頼性の低い誤った情報である可能性がある。このため、

レビューを参考にしたユーザが誤った判断を行うリスクがある。

また様々な意見が投稿されることにより、重要な意見が埋もれてしまう可能性もある。例えばある商品に対する数十件以上のレビューの中から、重要で価値のある意見を探し出す労力が必要であり、信頼のおける情報を取得するためには、そのレビュー投稿者（レビューア）の信頼性も考慮しなくてはならない。

このような課題に対処する手法として、本稿では金融工学で用いられるポートフォリオ理論の考えを利用したレビューポートフォリオを提案する。

ポートフォリオとは証券商品の組み合わせであり、ポートフォリオ理論は最適な証券の組み合わせを選択することを目的とする。このポートフォリオの性質をレビュー選択に利用したものがレビューポートフォリ

オである。レビューポートフォリオでは複数のレビューの中から信頼できる有用なレビューの組み合わせを提示する。その際レビューの組には一定の基準により対照的な性質を持つ組を選択する。これによりレビューの情報を利用する際のリスクを軽減することが本手法の目的となる。このような組み合わせ方はアドホックな形では既に Amazon 等のレビューサイトでも行われているが、本手法では統計的な分析を利用することで、対照的で有用なレビューを発見することを目的としている。

本稿では 2 章ではレビューデータに関する先行研究およびポートフォリオ理論の概略を紹介する。3 章では研究対象としている価格.com のレビューデータの傾向を調査した結果を示す。4 章では提案手法について述べる。5 章では提案手法に対して行った評価実験結果について記す。6 章では今後取り組む課題について考察する。

2. 関連研究

2.1 レビューデータを対象とした先行研究

本節ではレビューデータを対象とした先行研究について述べる。

小倉ら[3]は、レビューに対して付与される評価情報（レビュー情報に対する投票）に着目し、そのレビュー情報の支持率と参考度を算出する手法を提案している。これによりレビューの品質を考慮した対象へのスコアリングを行っている。

鈴木ら[4]はレビュー文中の評価表現を抽出する手法を提案している。評価表現を肯定意見・否定意見に分類する際に、semi-supervised な学習手法を用いることにより、分類性能の向上を図っている。

小林ら[5]はレビュー情報の信憑性を向上させる手法として、レビューのバイアスによる補正を提案している。彼らは我々が無意識のうちに影響をうけるバイアスがあることを確認し、これを用いてレビューの補正を行う手法を検討している。

これらレビューを対象とした研究課題の 1 つは、個人の投稿により発生するレビュー情報から、いかに有益な情報を取り出すかということである。様々なユーザからの意見が投稿される一方、誤った知識をもったユーザからの意見や、根拠のない批判意見なども投稿される場合がある。こういった多数のレビュー情報の中から信頼性の高い情報を発見することが、レビューを対象とした研究の目的の 1 つとして挙げられる。

2.2 ポートフォリオ理論

本節では提案手法で利用するポートフォリオ理論

の考え方について簡略に記す。

現代の資産運用において重要となるのが分散投資である。一般的に株式や債券などの証券に投資を行う際、特定の 1 証券に対して投資するのではなく、複数の証券に対して資産を分配して投資する。これは損失を起こすリスクを軽減しつつリターンを得るためである。このような複数の商品に対して資産を分散して投資を行うことを分散投資と呼び、分散投資を行った結果構成される資産構成をポートフォリオと呼ぶ。ポートフォリオ理論とはこれらポートフォリオの構成モデルを表す方法論であり、これにより構成内容を定量的に評価することが可能となる。

分散投資を行う際、重要な課題となるのがどの証券を選ぶか、そして選んだ証券にどれくらいの配分で資産を投資するかである。ポートフォリオ理論では、これら分散投資による資産運用を①損失を出してしまうリスク②投資により得られるリターンという 2 つの指標を用いる。そして過去の値動きをもとに、リスクを値動きの分散、リターンを収益の期待値と規定することで、証券の組み合わせを定量的に評価する。リスク・リターンを定量的に評価することで、分散投資の課題に対する指針を提供する。

ポートフォリオ理論に関する詳細は文献[6]、[7]および[8]に示されている。ここでは特に損失を出すリスクを最小とする手法について述べる。

ある 2 証券に対して分散投資を行うことを考える。これら 2 証券を運用する際リスクが最小となるのは、2 証券間に完全な負の相関があり（相関係数が-1）、これらに対して等量の投資を行う場合である。これは一方の値動きが他方の値動きを相殺するためであり、直感的にも明らかである。実際には、完全な負の相関をもつ証券モデルを構成するには、将来の値動きを完全に予測できないことから不可能であるが、証券市場の緩やかな成長を仮定した場合、最小のリスクでリターンを得ることが可能である。

本稿ではこれらポートフォリオ理論の概念をレビューの選択の場面で利用し、レビューを受け取った際に誤った判断をするリスクの軽減を試みる。

3. レビューデータの性質調査

レビューの分類・選択を統計的な観点から行う場合、その全体の性質や傾向を確認することが必要となる。本章では本稿で研究の対象とした価格.com のデータを元に行ったレビューデータの調査結果を示す。

3.1 研究対象

まず、調査対象となる価格.com に投稿されるレビュー例を図 1 に示す。



図 1：価格.com のレビュー例

価格.com ではレビューは商品の総合的な評価を表す「満足度」と「個別の評価項目」、商品への「コメント」から評価を行う。評価値は全て 1 から 5 の 5 段階評価である。

今回調査の対象としたのは価格.com の商品カテゴリ「デジタルカメラ」に属する 10623 件のレビューデータである。カテゴリ「デジタルカメラ」に属する商品は個別の評価項目として「デザイン」「画質」「操作性」「バッテリー」「携帯性」「機能性」「液晶」「ホールド感」の 8 つの項目に評価を行うことができる。

また各レビューには参考になったかどうかをユーザが投票することができる項目があり、「参考になった人数 / 評価者数」の形式で表される。本稿ではこれをレビューの支持率と呼び、以下の提案手法などに利用する。

3.2 評価値の分布調査

ここで上記レビューデータ 10623 件を対象にこれらの評価値を集計した結果である、各評価値の度数分布を示す。図 2 は総合評価である満足度の分布である。また図 3 は個別評価の 1 つ「携帯性」への評価値の分布である。なお 10623 件の中には、満足度以外の各個別評価が行われていない(無評価)レビューがあるが、これはデータの欠損値として扱い、分布の中には含まない。

図 2, 3 から全体的に最高の評価となる 5 を付けるレビューが多いことがわかる。総合評価、個別評価ともに高評価である 4 及び 5 を付けるレビューが全体の 80% から 90% にのぼる。逆に中評価、低評価を付けるレビューが非常に少ない。携帯性以外の個別評価も同様の分布をとっている。このような高評価に偏った傾

向は他の商品カテゴリにも見られる傾向である。

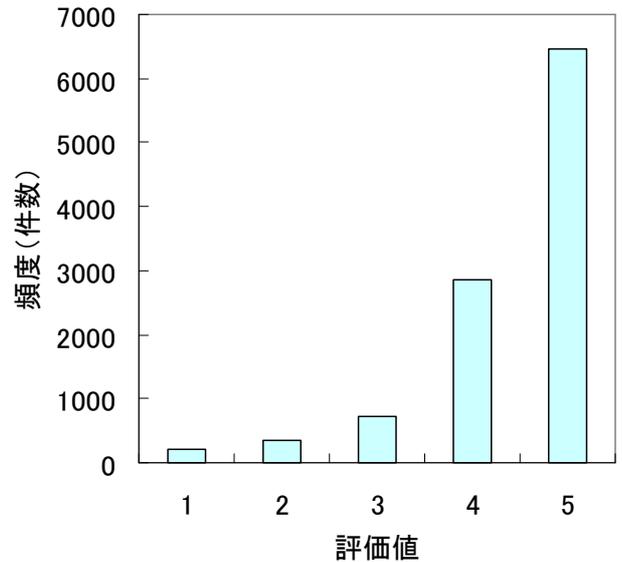


図 2：レビュー全体の「満足度」の評価値の分布

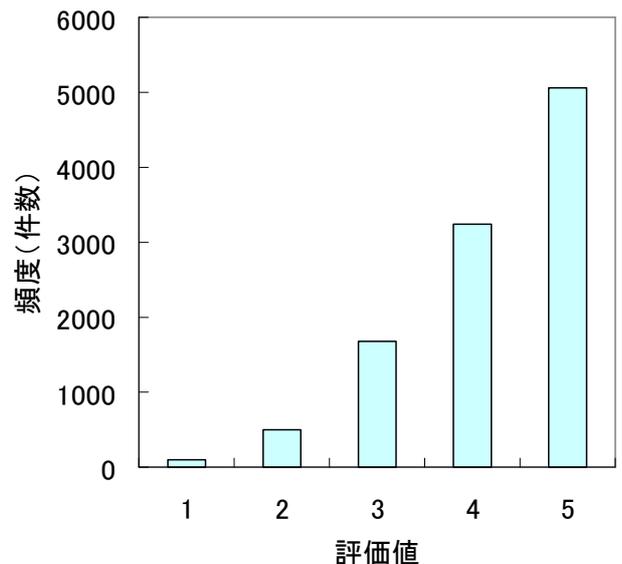


図 3：レビュー全体の「携帯性」の評価値の分布 (47 件が無評価であり分布から除外)

全体の傾向としては上記の結果が得られたが、ここで各商品カテゴリに対して深い見識をもつと考えられる“熟練者”を投稿したレビューの中から新たに定義し、その分布の傾向を調査した。熟練者の定義には、価格.com が提供するレビューランキングを用いる。これは各レビューのその商品カテゴリへのレビュー投稿件数とレビューの支持率を元に価格.com が独自に集計しているランキングである。デジタルカメラのランキング上位 30 位に入るレビューを熟練者と定義する。

彼らが投稿した 425 件のレビューに対して全体と同様に各評価値の度数分布を算出した。またこれらとの比較を行うため、熟練者レビュー以外の一般レビューからも同数の 425 件のレビューを無作為に抽出して比較を行った。評価項目「満足度」および「携帯性」の分布の比較を図 4、5 に示す。

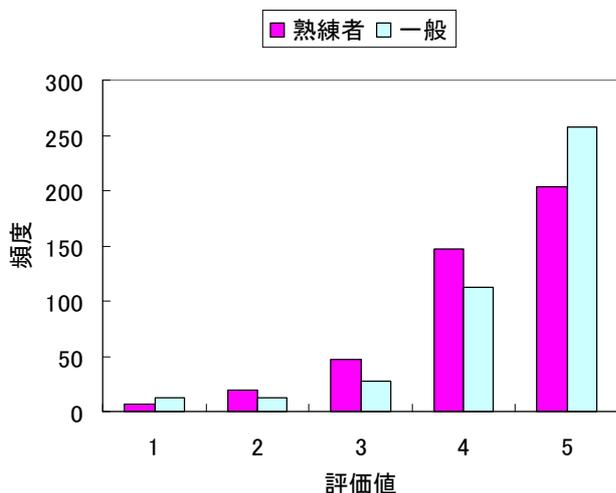


図 4：「満足度」の一般と熟練者との比較

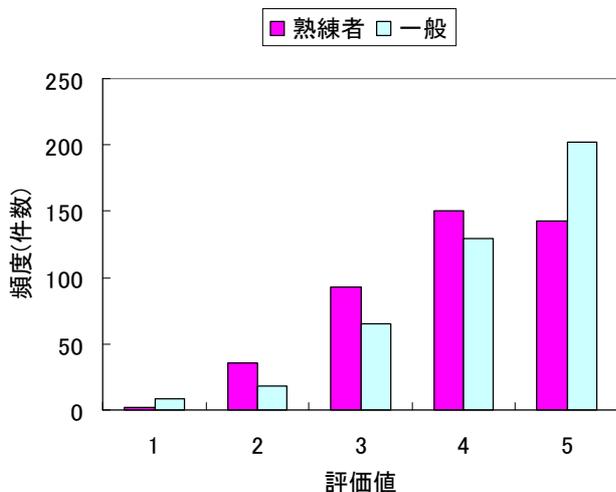


図 5：「携帯性」の一般と熟練者との比較
(熟練者は 2 件、一般は 2 件がそれぞれ無評価であるため分布から除外)

図 4 では熟練者の方が満足度に関する評価として 4 を付ける傾向が高いことがわかる。図 5 に示した携帯性の熟練者による評価ではそれが顕著に現れており、5 と評価した熟練者よりも頻度が高くなっている。このように、熟練者の評価傾向の分布は全体の傾向とは大きく異なることが確認された。

ここでより定量的な違いを示すために、一般と熟練者の各分布のエントロピーを求める。全体の中で、各

評価値の占める割合を確率で表現し一般と熟練者の情報量の違いを示す。

全体 425 件に対して評価値 i の占める割合を一般、熟練者それぞれ $P(\text{一般}_i)$ 、 $P(\text{熟}_i)$ とおくと、エントロピー $H(\text{一般}_i)$ 、 $H(\text{熟}_i)$ は以下のように表せる。対数関数の底は 2 とし、式中では省略する。

$$H(\text{一般}_i) = -\sum_{k=0}^5 P(\text{一般}_k) \log P(\text{一般}_k) \quad (3.1)$$

$$H(\text{熟}_i) = -\sum_{k=0}^5 P(\text{熟}_k) \log P(\text{熟}_k) \quad (3.2)$$

なお全体の分布と同様に各データ内には評価値が無評価のレビューも存在する。エントロピーの計算にはこれらの分布も含み評価 0 とおく。上記式中で評価値 i が 0 を考慮していることは無評価のレビューも存在するためである。

このような計算方法により各評価値の分布のエントロピーを、一般レビューと熟練者レビューで比較を行った結果を表 1 に示す。

表 1：熟練者と一般のエントロピーの比較

	エントロピー(bit)	
	熟練者	一般
デザイン	1.71	1.56
画質	1.89	1.75
操作性	1.85	1.84
バッテリー	2.20	1.99
携帯性	1.91	1.79
機能性	1.79	1.73
液晶	2.07	1.80
ホールド感	1.97	1.83
満足度	1.69	1.51

表 1 からわかるように、すべての評価項目で熟練者のエントロピーが一般レビューを上回っている。熟練者は傾向として評価のバラつきが見られるため、相対的にエントロピーも増加することになる。

本稿では、全体的に最高評価を付ける割合が多い中、何らかの理由で最高評価をつける割合が小さく、傾向として評価が分散する熟練者を、他のレビュー以上の情報を持つと推定する。提案手法ではこれら熟練者のレビューを後述する有用なレビューの選択の際の基準

の1つとして採用している。

4. 提案手法

本章では提案手法であるレビューポートフォリオについて述べる。ある商品に対して複数のレビューが存在する場合、レビューの数が多ければ全てを確認することは困難である。この問題を解決するため、本手法では信頼性が高く、なおかつ参考となる有用なレビューを提示する。その際、第2章で記した証券の例のように特定の1レビューのみを参照することは誤った判断を招く可能性がある。信頼できるレビューを選択できたとしても、商品を多面的な観点から判断するためには1つのレビューでは情報が不足することが考えられるためである。

そこで提案手法であるレビューポートフォリオでは、ユーザにとって閲覧が負担とならない数のレビューを複数提示する。この際提示されるレビューの組は購買決定時に誤判断をするリスクを軽減するために、レビューに対する統計的分析をもとに、性質の異なる組み合わせを選択する。例えば高評価のレビューと低評価のレビューの組み合わせのように評価値が全く異なるレビューを合わせて提示することで、商品に関する多面的な意見をユーザに提示する。このようにして選択されたレビューの組には、その商品への良い面、悪い面の双方に関する意見が記述されていると期待できる。そのため誤判断のリスクを軽減すると同時に、商品の多様な意見を取り入れることができる。

4.1 提案手法の構成

ここで図6に提案手法の構成を示す。

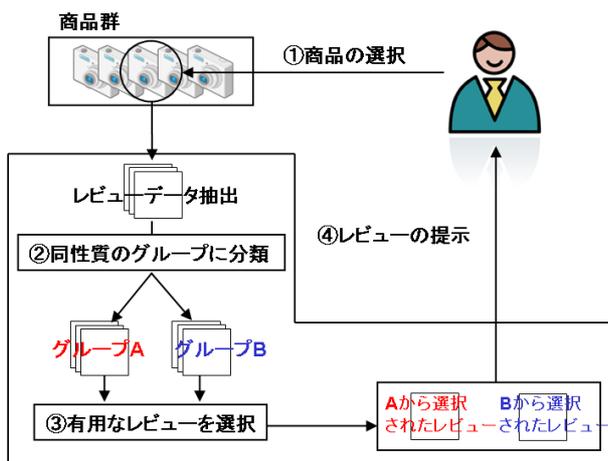


図6：提案手法によるシステム構成図

図6に示すように、まず複数のレビューを一定の基準の元で分類する。図6では2つのグループに分類を行っているが、2つ以上のグループへの分類が行われ

てもよい。分類には様々な手法が考えられるが、今回は階層的クラスタリングを用いて4~6つのクラスターを構成し、これらを分類されたグループとする。詳細については4.2節で詳しく解説する。

次に分類によって分けられた同性質をもつレビューのグループの中から有用なレビューを選択する。有用なレビューの選択に関しては様々な基準が考えられるが、本稿では前章の傾向調査の結果と直感的な方法を利用したルールを作成する。このルールに従う任意のレビューを選択する。より客観的・定量的な手法の検討は今後の課題とする。選択手法については4.3節で述べる。

前述のようにAmazonでは本手法と同様のサービスが行われている。Amazonでは商品に対する点数(1から5)と、レビューに対する支持率をもとに、より多くの人々から支持されている高評価と低評価のレビューを同時提示する方法をとっている。本手法ではレビューの分類と選択を統計的な分析・手法に基づき行うことでより有用なレビューの発見を目指す。

4.2 同性質のグループへの分類

ここで、提案手法におけるレビューの分類について記す。レビューの分類には様々な手法が考えられるが、今回は各レビューの評価値(操作性、機能性、満足度など)をもとに階層的クラスタリングを行う。その結果得られたクラスターにより、レビューの分類を行う。

具体的な手法としては最遠傍法を用いる。実際の商品「LUMIX DMC-LX2」に対して行った結果得られたデンドログラムを図7に示す。

各データのラベルには、投稿者であるレビューの名前がついている。このデンドログラムから4つから6つ程度のクラスターを構成するように各データをグループ化して分類を行う。

4.3 有用なレビューの選択

前節の階層的クラスタリングを用いて構成した各クラスターから有用と考えられるレビューの選択を行う。選択にはレビューデータから得られる以下の情報を用いる。

- ・ レビューのコメント文の文字数
- ・ レビューの支持率
- ・ 各評価値間の分散

各評価値間の分散とは、付与される9つの評価項目の分散をレビュー毎に算出した値である。

上記の情報を用いて以下に示すような選択ルールを作成した。

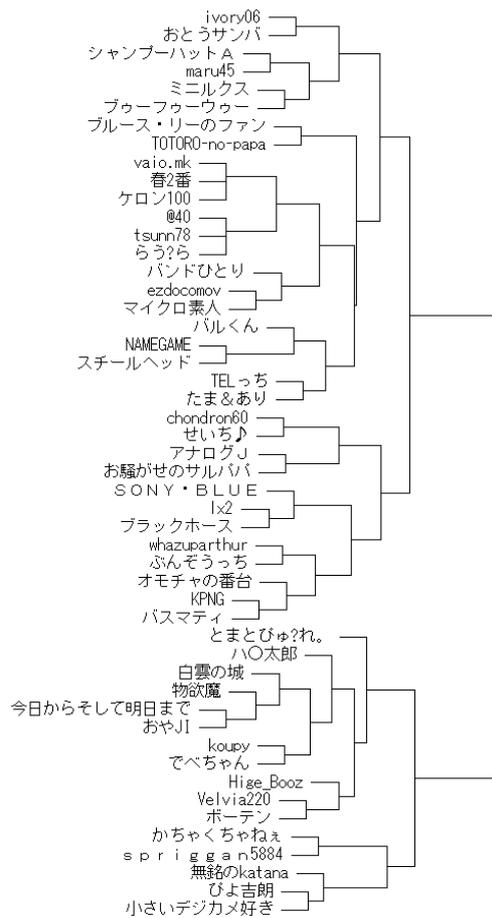


図 7：最遠傍法によるクラスタリング結果

選択ルール

各クラスタ内から 1 件ずつ選択する

- ① クラスタ内に熟練者が 1 人のみ投稿されている場合
⇒ 投稿した熟練者のレビューを選択する
- ② クラスタ内に熟練者が 1 人も投稿していない場合
⇒ 以下のルールを満たすものを選択する
 - 商品内のコメント文字数の平均値を超えている
 - 投票が 5 件以上で支持率が 20% 以上である
 - クラスタ内での評価値間の分散が相対的に大きい
 これらを満たすレビューの中で、ランダムに選択する
- ③ クラスタ内に熟練者が複数投稿した場合
⇒ ②と同様のルールを熟練者のみを対象に適用して選択する。

文字数に閾値を設けたのは、多様な意見をもつレビューには一定の情報量が必要であると考えられるためである。また評価値間の分散が大きい場合、評価にバラつきがあることになる。クラスタの中で評価にバラつきがあるレビューの方が良い面と悪い面の両方を記述している可能性が高い。各クラスタ内で分散が極端に小さいレビューに関してはそのクラスタから除外する。

このようなルールに従って得られた集合を有用なレビューの組として提示する。

5. 評価実験

本章では、提案手法に対して行った評価実験の結果を示す。評価実験では選択されたレビュー情報の多様性に着目する。

本手法の目的の 1 つは異なる意見をもつレビューの提示が目的であるため、選択されたレビュー群が各商品の肯定的な意見と否定的な意見を網羅的に記述されていることが望ましいと考えられる。そこで、提案手法を用いて選択したデータと一般的なデータを比較する。

5.1 使用データ

価格.com の商品カテゴリ、デジタルカメラに属する商品の中で、コメント付きのレビューが 50 件以上投稿されている 8 商品を選択する。各商品に投稿されたレビューデータを 1 つの集合とみなし、4 章で示した提案手法を用いて分析する。得られたレビュー群を“評価データ”と呼ぶ。

また各商品に投稿されるレビューのコメント文の平均をとり、評価データとして選択されなかったレビューの中から文字数が各商品内の平均値に近い順に選択する。これを提案手法の評価データと比較する“比較データ”と呼ぶ。

5.2 評価方法

得られた評価データと比較データのコメント文に着目する。コメント文を読み、その記述の中でカメラのレビューに用いられる 8 つの個別評価項目（デザイン、操作性など）に関する記述が行われているかを確認する。次にその記述が肯定的な意見または否定的な意見のどちらを言及しているかを判定する。さらに選択されたレビュー全てを合わせた意見の言及数と言及率を確認する。8 つの項目それぞれに肯定・否定の意見を判定するため、全 16 項目からなる。これらの調査を評価データおよび比較データそれぞれに行い、言及数と言及率を比較する。

言及された文かどうかは、明らかにその項目への肯定・否定意見が記述されているものに限定する。例えば「～の画質は素晴らしい」や「～手振れ補正に関しては良好です」などを肯定意見、「～ノイズがひどいため良くない」や「～質感がイマイチ」などを否定意見とする。このほかの「～に関しては問題ない」や「～は悪くはない」などの肯定・否定のどちらともとれない表現やどちらとも判断できる表現に関しては、考慮しないこととする。また評価者のカメラに関する知識ではどの評価項目に関する記述か判断できない記述も考慮しない。ひとつのレビューの中で複数回言及が行われた場合も、1回の言及と数える。

5.3 実験結果

ここでは実験結果を示す。まず8商品に対して行った全体の言及率を表2に示す。

表2：8商品の言及率の比較

	選択レビュー数	言及率(言及された項目/全項目)	
		評価データ	比較データ
商品1	5件	69%(11/16)	69%(11/16)
商品2	5件	56%(9/16)	25%(4/16)
商品3	5件	56%(9/16)	50%(8/16)
商品4	5件	81%(13/16)	69%(11/16)
商品5	6件	63%(10/16)	69%(11/16)
商品6	4件	25%(4/16)	25%(4/16)
商品7	5件	31%(5/16)	44%(7/16)
商品8	4件	50%(8/16)	50%(8/16)

このように、8商品中3商品が比較データと比べて、言及率が高いことがわかる。言及率が同じになったのは3商品であり、比較データよりも低くなったのは2商品であった。

表3：商品2における評価データの言及数詳細

	言及の有無	
	肯定意見	否定意見
デザイン	○	○
画質	○○○	×
操作性	×	×
バッテリー	○	×
携帯性	×	×
機能性	○	○
液晶	○○	×
ホールド感	○	○
全言及回数	12回	

表4：商品2における比較データの言及数詳細

	言及の有無	
	肯定意見	否定意見
デザイン	×	×
画質	○○○○○	○
操作性	×	×
バッテリー	○○	×
携帯性	×	×
機能性	×	×
液晶	×	×
ホールド感	×	○
全言及回数	9回	

次に評価データと比較データの言及率に大きな差が生じた商品2の評価結果の詳細を表3及び表4に示す。言及が行われた項目には○を、言及されなかった項目には×を記す。また言及された項目では、その言及の回数を○の数で示すこととする。

表3および表4の結果から言及率が56%と25%であることに比べ、言及された回数の総和の差は12回と9回で小さいことがわかる。

5.4 考察

表2に示したように、評価データが比較データの言及率を上回る商品がある反面、比較データの方が高い言及率になる商品もあった。言及が明確にされたかどうかはレビューの質を考慮しただけでは、判断が難しいことが課題として挙げられる。有用であると選択したデータの中には明確な批評意見の記述のないレビューもあった。

また今回レビュー選択の際に熟練者のレビューを優先的に選択するルールを用いた。しかし熟練者のレビューの中には支持率が低く、情報量に影響するコメントの文字数も少ないレビューもあった。全体的には、熟練者は他にはない貴重な意見をもつと推定できるが、中にはノイズの多い意見を投稿するレビューも見られた。今回は評価値の分布の差異があることから熟練者を選定したが、別の基準に基づいて熟練者を発見することが重要ではないかと考えられる。熟練者、つまり他のレビュー以上の知識を持つレビューをレビューデータから発見することは、有用な情報を発見することにもつながり、今後の課題である。

全体の結果とは別に表3、表4で示した結果では、言及率の大きな差と比較して、明確な言及がなされた回数の差が小さいことがわかる。比較データの場合画質の肯定意見の言及が5回行われており、意見に偏りが生まれている。それに対して提案手法を用いてレビ

ューを選択した場合、意見が全体的に分散していることが確認できた。この商品2は本稿で提案した手法の目的が達成されている例であるといえる。

6. おわりに

本稿では、提案手法であるレビューポートフォリオの提案を行い、その評価実験結果を示した。レビューデータの参照に伴うリスクの軽減とより多様な意見の参照を目的としたが、本章では今後の課題について述べる。

まず本手法におけるレビューの分類と選択には、アドホックな処理が多いことが課題としてあげられる。レビューの分類に用いた階層的クラスタリングのクラスタ決定の際は、ある階層におけるクラスタを用いることが統計的に最適なものかどうかの判定を行っていない。これには、モデル選択の場面などで利用される赤池情報量基準や最小記述長といった基準がクラスタを選択する際の判定基準に利用できると考えられる。

また分類されたレビューの中から有用なレビューの選択を行ったが、明確な裏付けのある基準を用いることが望ましい。5.4節で記したように熟練者の定義には、より明確な分析に基づいて選定を行う必要があると考えられる。今回の評価実験のように、選択されたレビュー群の多様性を重要視するのであれば、レビュー文内に明確な評価表現があるかどうかでも有用なレビューであることの基準にする必要がある。

選択の場面における今後の新たなアプローチとしては、レビューのコメント文へのテキスト解析を行うことが挙げられる。有用なレビューかどうかを今回のような周辺情報から判断することは難しく、実際にはそのレビューのコメント文の品質によるところが大きい。評価表現の抽出など様々な分析が必要ではないかと考えられる。またレビュー1件の中でも重要な記述とノイズに近い記述の両方が含まれていることがある。レビューをそのまま提示するのではなく、新たな手法を用いてノイズに近い記述を除去するなどの手法も有効であると考えられる。

コメント文を利用した解析手法の1つとして、レビューデータの生成モデルを構築することも検討すべき手法として挙げられる。レビューを特徴付ける記述などが、どのような過程を経て生成されるかを表現することができれば、レビュー全体の解析に役立てることができると考えられる。

このように今回提案したレビューポートフォリオには今後さまざまなアプローチから改善の余地があると考えられる。有用なレビューの提示によるリスク軽減を行うための枠組みについてさらに検討していく方針である。

謝 辞

本研究の一部は文部科学省私立大学戦略的研究基盤形成支援事業「芸術・文化分野の資料デジタル化と活用を軸とした研究資源共有化研究」の支援を受けている。

文 献

- [1] Amazon.co.jp, <http://www.amazon.co.jp/>
- [2] 価格.com, <http://kakaku.com/>
- [3] 小倉達矢, 宍戸開, 今藤紀子, 山口実靖, 浅谷耕一, “レビューサイトにおける良質なレビューの特性とそれを考慮した評判情報の抽出に関する一考察”, 第19回データ工学ワークショップ(DEWS2008) B8-5, 2008.
- [4] 鈴木 泰裕, 高村 大也, 奥村 学, “Weblogを対象とした評価表現抽出”, 人工知能学会研究資料SIG-SW&ONT-A401-02, 2004.
- [5] 小林卓弥, 大島裕明, 小山聡, 田中克己, “レビューアーのプロファイルと地域性に起因するバイアス補正に基づくレビュー情報の信憑性向上”, 第19回データ工学ワークショップ(DEWS2008) B8-4, 2008.
- [6] 藤田岳彦, “道具としての金融工学”, 日本実業出版社, 2005.
- [7] 野口悠紀雄, 藤井真理子, “金融工学—ポートフォリオ選択と派生資産の経済分析—”, ダイアモンド社, 2000.
- [8] 今野浩, “理財工学—平均・分散モデルとその拡張—”, 日科技連出版社, 1995.
- [9] Jiawei Han, Micheline Kamber, “Data Mining: Concepts and Techniques Second Edition”, Morgan Kaufmann, 2006.
- [10] 加藤直樹, 羽室行信, 矢田勝俊, “シリーズオペレーションズ・リサーチ 2 データマイニングとその応用”, 朝倉書店, 2008.
- [11] 平田廣則 “情報理論のエッセンス”, 昭晃堂, 2003.