

テキストマイニングを用いた口コミ分析による点数評価の信頼性確認手法

Reliability confirmation method of score by personal communication using text mining

谷口 佑子^{*1}
Yuko TANIGUCHI

津田 和彦^{*1}
Kazuhiko TSUDA

^{*1} 筑波大学大学院ビジネス科学研究科
Graduate School of Business Sciences, University of TSUKUBA

The product evaluation written by customers on the website is important information influencing the customer's purchase consciousness. In many cases, the evaluation page is composing of score for the predetermined items and personal communications. However, the difference in the score is very small and it is hard to say the result correctly evaluated the product. Therefore, in this paper, we propose a method to confirm the reliability of score by comparing the result extracted from score and personal communication using text mining and data on the product evaluation page of the camera.

1. はじめに

インターネットの普及に伴いネットショッピングの世帯利用率は伸び続けている[内閣府 2015][総務省 2015]. 家電製品においても、ネットショッピングの利用は伸びつつあり、今日ではネットショッピング全体の 10%を家電製品が占めている[統計局 2015]. また、家電製品をネットにて購入する人数は、2012 年からの 8 年間で 1.8 倍を超えると予想されている[日本政策投資銀行 2013]. かつて、家電製品の購入検討者は、店頭へ出向き、店員やメーカーの専門家にアドバイスをもらいながら購入する製品の選択していた。しかし、今日では、多くの購入検討者がネット上に存在する購入候補となる製品の情報を収集し比較する事で、購入する製品を決めている。

製品に対する情報として有用と言われ、多くの購入検討者が活用しているのが、当該商品の使用者が書き込んだ製品評価ページの情報である。製品評価ページは、主に点数評価と口コミ情報から構成されている。口コミ情報は、そのまま列挙されているだけであり、多くの口コミが集められているサイトでは、全ての口コミを読むことは困難である。一方、点数評価は、平均とか点数の分布など整理された情報が示されている。けれども、その情報の信頼性に疑問を持つ購入検討者は多い。

そこで本研究では、カメラを購入した顧客から寄せられた点数評価と製品への口コミ情報を用いて、点数評価の信頼性確認手法を検討する。具体的には、点数評価である項目と、それに付随する口コミを分析することで、評価レビューにおける信頼性の確認を導くことを目的としている。

2. 製品購入サイトのレビュー情報

点数評価と口コミ情報の関連性を説いた研究は、日本語の宿泊レビューに関するものがあつた[辻井 2012]. しかし、この信頼性が英語について導かれているものではなく、またカメラについて書かれている先行研究も存在しない。

多くの欧米の家電メーカーの製品購入ページには、製品のレビュー情報が掲載されている。製品購入ページに記載された点数評価は 5 または 4 を付けられることが多い。当該製品を気に入る、購入した人の評価なのだから、ある意味当然の評価である。

製品のレビュー情報における点数評価には、図 1 に示すよ

うに、総合評価以外に、製品の機能や使い勝手など製品の特長についての評価項目を持っている場合が多い。

本研究では、総合評価だけではなく、これらの特徴について評価する項目にも着目し、その点数評価の信頼性についても確認できる手法を提案する。



図 1: 家電製品におけるレビュー

3. 点数評価の結果

表 1 に、Sony の製品サイトからレンズ交換が可能なカメラ Interchangeable Lens Camera (<http://www.sony.com/electronics/interchangeable-lens-cameras/>)を対象に、顧客が総合評価に付与した点数の集計結果を示す。

結果からわかるように、全体の 66.79%が最も評価が良い 5 を示しており、概ね良いとされる 4 を合わせると 95%に近い数字になる。一方で、評価が悪い 1 および 2 は全体の 2.34%しか存在していない。つまり、点数評価の結果は非常に大きな偏りがあることがわかる。

Score	Count	Count/Total
5	939	66.79%
4	389	27.67%
3	45	3.20%
2	13	0.92%
1	20	1.42%
Total	1,406	

表 1: レンズ交換可能なカメラの点数評価

次に、特徴を評価するために準備されている Picture Quality(画質), Features(機能), Design(デザイン), Ease of Use(使い勝手)の 4 項目の点数評価について集計した。表 2 は、総合評価を行った全ての人を対象とした各項目の点数評価を集計した結果であり、表 3 は、総合評価で 5 を付けた人の各項目の点数評価を集計した結果である。

Score	PICTURE QUALITY	FEATURES	DESIGN	EASE OF USE
5	1,075	937	865	653
4	271	397	458	579
3	28	34	59	126
2	9	14	11	21
1	10	10	7	20
0	13	14	5	7
Total	1,406	1,406	1,405	1,406
5	76.46%	66.64%	61.52%	46.44%
4	19.27%	28.24%	32.57%	41.18%
3	1.99%	2.42%	4.20%	8.96%
2	0.64%	1.00%	0.78%	1.49%
1	0.71%	0.71%	0.50%	1.42%
0	0.92%	1.00%	0.36%	0.50%

表 2: 各項目の点数評価

Score	PICTURE QUALITY	FEATURES	DESIGN	EASE OF USE
5	848	773	742	580
4	78	150	186	317
3	2	3	6	29
2	1	1	1	5
1	0	2	0	3
0	10	10	3	5
Total	939	939	938	939
5	90.31%	82.32%	79.02%	61.77%
4	8.31%	15.97%	19.81%	33.76%
3	0.21%	0.32%	0.64%	3.09%
2	0.11%	0.11%	0.11%	0.53%
1	0.00%	0.21%	0.00%	0.32%
0	1.06%	1.06%	0.32%	0.53%

表 3: Score5 の人の各項目内訳

これらのデータからわかるように、総合評価が 5 だとしても、全ての各項目において高い評価が得られるわけではない、ということが見える。例えば、表 3 において Picture Quality, Features や Design は評価が高いのに対して、Ease of Use には不満を持っている人が 40% 近くいることがわかる。この結果より、総合評価では満点を与えていても、必ずしも全ての事項で満足しているというわけではないと言える。

このことを拡大解釈すると、総合評価として高い評価であっても不満はあり、その不満点はここで評価した 4 項目以外である可能性も否めない。さらには、不満点があっても総合評価として高い評価を与えているという事は、ここで評価した 4 項目以外で、吐出して満足な点が存在する可能性もある。

4. 点数評価の信頼性確認手法

4.1 点数評価の信頼性

点数評価の信頼性を確認するため、口コミをテキストマイニングにて分析する。テキストマイニングでは、Positive, Neutral, および Negative な表現を抽出する感性評価を実施した。具体的には、総合評価の結果の各点数の口コミから、各感性の出現頻度をカウントした。その結果を表 5 に示す。なお、本分析で利用したテキストマイニングツールは、NTT データ数理システムのテキストマイニングスタジオ 6.0.1 で、辞書などの追加は行わず初期設定の状態で使用した。

Score	5	4	3	2	1
Positive	497	4	29	9	2
Neutral	442	363	9	3	16
Negative	0	21	13	4	2
Total	939	389	45	13	20
Positive	52.93%	1.03%	64.44%	69.23%	10.00%
Neutral	47.07%	93.32%	20.00%	23.08%	80.00%
Negative	0.00%	5.40%	28.89%	30.77%	10.00%

表 4: 総合評価と口コミ分析の感性評価

その結果、総合評価 4 では Neutral に分類される文章が最も多く 90% を超えることが分かった。また、総合評価 5 の場合でも

口コミの約半数は Neutral を示すことが分かった。また、高評価では、Negative な口コミが少ないことも分かった。

4.2 テキストマイニングによる情報抽出

テキストマイニングによる分析では、その他にも様々な情報を獲得することができる。ここでは、総合評価が 4 で Neutral に評価された 363 件の口コミに対してテキストマイニングを実施した。その際、削除単語 3、同義語 2 語を登録した。テキストマイニングにより、感性情報を有する単語と係り受け関係となった単語の頻度ランキングの TOP10 を表 5 に示す。

Noun	Count	Relations	Count
Photo	117	Photo-Take	34
Lense	114	Lense-Use	14
Quality	114	Lense-Buy	8
Feature	83	Lense-Change	6
Image	75	Lense-Have	6
Excellent	61	Canon-Use	5
Button	43	Feature-Have	5
Menu	41	Problem-Have	5
Time	40	Quality-Good	5
Video	40	Shot-Take	5

表 5: テキストマイニング結果

Neutral に分類された文章の中では、レンズの品質に関するコメントが多く書かれていることがわかった。また、レンズに対しては、使用、購入、交換など様々な意見が述べられていることが見えてきた。このことから、口コミ情報をテキストマイニングにて分析すれば、顧客が何について興味を持ち、どう評価しているのかという様々な情報を入手する事が可能であることが分かった。

5. おわりに

本研究では、製品購入サイトの点数評価と顧客の感想であるコメントを用いて、点数評価とコメントの信頼性の確認手法を提案した。

その結果、表 4 に示す総合評価の各点数における Positive, Neutral, および Negative な口コミの表現数から、総合評価は妥当な点数であると言えることを確認した。さらには、総合得点が高い評価であっても、全ての項目に対して満足している訳でない、という事もわかった。

今後の取り組みとして、顧客が何をどう評価して当該製品を購入したのかという事と、当該製品の何に不満を持っているのかという事を把握することに取り組む。さらには抽出した情報を用いて製品改善に向けた指標を示すシステムを構築する事を検討する。

参考文献

- [総務省 2015] 総務省:情報通信白書 インターネットショッピングの利用状況(2015)
- [統計局 2015] 総務省統計局:家計消費状況調査年報 インターネットを利用した支出の状況(2015)
- [内閣府 2015] 内閣府: 経済社会総合研究所 消費動向調査(2015)
- [日本政策投資銀行 2013] 株式会社日本政策投資銀行 産業調査部 Eコマース市場の拡大と小売業への影響(2013)
- [辻井 2012] 辻井 康一, 津田 和彦, テキストマイニングを用いた宿泊レビューからの注目情報抽出方法, 情報処理学会デジタルプラクティス(2012)