

製品の顧客評価による Web ショッピングの売上向上戦略

日大生産工(院) ○松本 光裕 日大生産工 豊谷 純

1. まえがき

日本の Web ショッピングの普及は 1995 年後である。インターネットの普及とともに、今日に至るまで Amazon や楽天等のインターネットショッピングモールや販売商品の増加、時間に関係なくいつでも商品が購入できる時代となっている。製品、商品の購入には、消費者が「星の数」や「カスタマーレビュー」等で商品の評価を行っている。これらによる評価が製品に対する顧客の信頼度やロイヤリティーの向上につながる。本稿は、Web ショッピングにおける顧客からの信頼度を向上させ、ロイヤリティーをつけるには、どのような文章が手法として効果的であるかを明らかにするために、分析手法としてテキストマイニングを利用する。

2. 購買心理との関連づけ

顧客がショッピングサイトにて商品を購入する段階に至るまでの行動について考えていきたい。購買心理の代表的なものとして「AIDMA の法則」³⁾と「AISAS の法則」がある。この 2 つの法則は「認知段階」「感情段階」「行動段階」のプロセスに分けられる。

2 つの法則の違いは「行動段階」のプロセスにある。AIDMA の行動段階では、その商品を欲しいという Desire (欲求) を引き出し、Memory (記憶) に焼付け、店頭での Action (購買) のプロセスを前提にする。一方で AISAS はインターネット普及以降の消費行動の特徴として、消費者は購入前に「Search (検索)」し、購入後に「Share (情報共有)」するという 2 つのプロセスを前提にしている。

Table.1 AIDMA モデル

AIDMA モデル	
認知段階	A : Attention (注 意)
感情段階	I : Interest (興味, 関心)
	D : Desire (欲求)
	M : Memory (記憶)
行動段階	A : Action (行動)

Table.2 AISAS モデル

AISAS モデル	
認知段階	A : Attention (注 意)
感情段階	I : Interest (興味, 関心)
行動段階	S : Search (検索)
	A : Action (行動)
	S : Share (情報共有)

3. 形態素解析の利用

3.1. 形態素解析とは⁴⁾

形態素とは、言語学で使われる専門用語で「意味の最小の単位」と言われている。それを文の形態素の単位に分割することを「形態素解析」という。

形態素解析の特徴として、文を形態素の単位に分割すると同時に、各形態素の品詞を特定することができる。

本稿における解析は、統計解析の代表的なものである「R」⁴⁾と形態素解析の 1 つである「MeCab」を利用した「RMeCab」⁴⁾、さらに形態素の関係性を表すために「KH Coder」⁵⁾も利用する。

3.2. RMeCab について

RMeCab は石田基広が独自に開発したパッケージであり、R から日本語の文章やファイルを指定して MeCab に解析させ、その結果を R で標準的なデータ形式に変

* The sales improvement strategy of Web shopping by customer evaluation of a product

換してデータ形式に変換するものである。パッケージとは、データやプログラムをまとめたものを意味する。

Web ショッピングの口コミやカスタマーレビュー等から寄せられるコメントをテキストファイルに保存し、フォルダとテキストファイルの場所を指定して読み込ませ、RMeCab で扱う関数を利用することで、テキストや文章を品詞ごとに分類、品詞の出現頻度を表すことが出来る。

3.3. KH Coder⁵⁾について

KH Coder とは、内容分析（計量テキスト分析）もしくはテキストマイニングのためのフリーソフトウェアである。立命館大学産業社会学部の樋口耕一が、新聞記事・質問紙調査における自由回答項目・インタビュー記録など、社会調査によって得られる様々なテキスト型データを計量的に分析するために制作されたものである。

4. 解析例

本論では Web ショッピングサイトにある口コミ・コメントの情報対象に解析を行う。初めに口コミやレビュー情報を収集するサイトとして「Amazon.com」を利用する。選択する製品・商品については、「MacBook Air」について口コミ、レビュー情報を集める。Amazon では「Mac Book Air」のレビュー情報を RMeCab による形態素解析を行い、形態素の「品詞」の抽出条件を指定する。

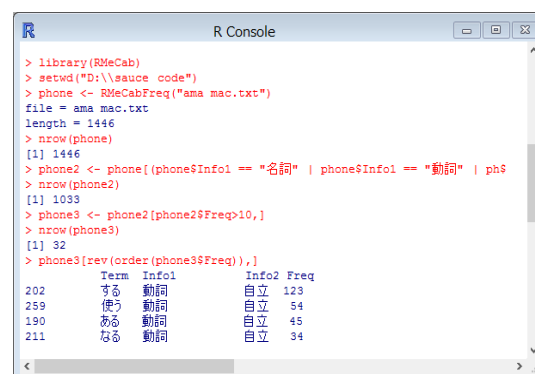
4.1. 実験対象とするサイト

Amazon.com

4.2. 製品または商品の選択

- ・ MacBook Air 1.7GHz Core i5/11.6
- ・ MacBook Air 1.7GHz Core i5/13.3

上記 2 つの製品の商品に対するコメントを収集し、1 つのデータとしてテキストファイル保存する。



```
> library(RMeCab)
> setwd("D:\\sauce code")
> phone <- RMeCabFreq("ama_mac.txt")
file = ama_mac.txt
length = 1446
> nrow(phone)
[1] 1446
> phone2 <- phone[(phone$Info1 == "名詞" | phone$Info1 == "動詞" | ph$
> nrow(phone2)
[1] 1033
> phone3 <- phone2[phone2$Freq>10,]
> nrow(phone3)
[1] 32
> phone3[rev(order(phone3$Freq)),]
  Term      Info1      Info2 Freq
202   する      動詞      自立    123
259   使う      動詞      自立     54
190   ある      動詞      自立     45
211   なる      動詞      自立     34
```

Fig. 1 RMeCab による入力結果

形態素の全体数	1446
条件による絞り込み	1033
頻度 10 回以上の形態素	32

Fig. 2 入力結果の概要

4.3. 口コミ・コメント数、抽出条件

4.3.1. 口コミ・コメント等の件数

件数：54 件

4.3.2. 抽出条件

- (1) 名詞，形容詞，動詞を抽出する
- (2) 非自立，数，接尾，自立を除く

5. 解析結果

5.1. RMeCab による形態素解析の結果

RMeCab の関数を利用し、保存したテキストファイルから形態素を抽出した入力結果が Fig. 1 である。4.3.2 の抽出条件を基に抽出された形態素は 1446 個であるが、全体を見通すにはまだ数が多い。軽量言語学とよばれる分野の研究によれば、テキストの中に現れる形態素の半数以上が 1 回しか出現しなく、残りの半分程度が出現回数 2 回だけの形態素である。そこで形態素の頻度が 1 回や 2 回だけではなく、形態素の頻度としてより多く使われている言語を抜き取るために、本稿では形態素の頻度 10 回以上を対象に絞り込みを行い、1000 以上ある形態素を減らすこととした。

5.2. KH Coder による形態素の抽出結果

KH Coder による形態素の抽出は、5.1 で利用した抽出条件 (4.3.2) を基に解析を行う。KH Coder の初期設定の関係により抽出方法を統一するために、「Chasen

「茶釜」から R と同様に「MeCab」に設定を変更する。また KH Coder では抽出条件 (4.3.2) に「形容動詞」は含まないが、RMeCab 上では「名詞」の「形容動詞語幹」に分類されるため、RMeCab と同一の結果として抽出する。

このように 5.1 の RMeCab による形態素解析の結果と KH Coder の形態素解析の同じ頻度 (出現回数) を抽出したものが Table.3 である。

Table.3 形態素の頻度数 (出現回数)

形態素	品詞	出現回数
ない・無い	形容詞	29
出来る・できる	動詞	26
購入・買う	サ変名詞・動詞	23
必要	形容動詞 (名詞, 形容動詞語幹)	14
動く	動詞	14
快適	形容動詞 (名詞, 形容動詞語幹)	14
スペック	名詞	14
使用	サ変名詞	12
使える	動詞	12
問題	ナイ形容	11
起動	サ変名詞	11
キーボード	名詞	11
満足	サ変名詞	10
作業	サ変名詞	10
軽い	形容詞	10

6. 製品の購入要因の調査

6.1. 形態素解析の頻度数

Table.3 から解析結果として表された形態素の頻度数をテキストファイル内にある各形態素を検索し、消費者の購入要因との関連性を調査した。

本稿では Table.3 の形態素の頻度数から、「購入・買う」「問題」「満足」の3つについて記す。

・要因 1. 購入・買う

「購入・買う」は、「薄さと軽さ」、「処理速度が速い」「へたなスマホ使うよりフルキーボードで使いやすい」等、「なぜこの商品を購入したか」についての理由を明確に探ることが出来る。また「購入」する人向けに「注意点」をレビューに記述する人もいる。

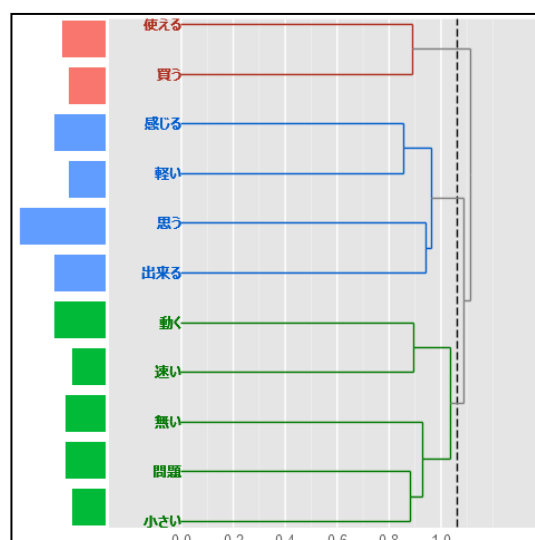


Fig.3 形態素の階層的クラスター分析

・要因 2. 問題

「問題」は、製品・商品における問題点として、「拡張性がない」「USB の電源容量が小さい」等のレビュー情報として挙げられるが、「問題ない」という否定的な回答として使われることが多く、人間が見極める必要がある。

・要因 3. 満足

製品・商品における重要な製品評価や購入要因を探ることが出来る形態素であると言える。主に「デザイン」「質感」「タッチパッド」「動作の速さ」「薄さ・形」等、各項目で評価がされており、結果として「満足」という文章が含まれていることが多い。

6.2. 形態素の関係性

6.2.1. 階層的クラスター分析⁶⁾

Fig.3 は、KH Coder を利用して階層的クラスター分析を行った。階層クラスター分析を行うとデンドログラム (樹形図) が出力され、各形態素がクラスターとして結合されていくかの過程が見られる。

本稿では各形態素の「関連性」を調査するために Table.3 に幾つか含まれている形態素をグループ化する。「小さい」と「問題」という 1 つのクラスターに「無い」が組み込まれ、「動く」と「速い」というもう一方のクラスターと統合され、1 つのグループとして分けられる。

最終的には「使える」と「買う」のクラスターでまとめられる。必ずしも隣接された語が同じクラスターに分類されるとは限らず、多次元尺度構成法とクラスター分析の2つの分析方法でデータを適応し、結果の違いを見極めることが出来る。

6.2.2. 多次元尺度構成法⁷⁾

多次元尺度構成法は、個体間の親近性データを、2次元あるいは3次元空間に類似したものを近く、そうでないものを遠くに配置する方法で、データの構造を考察する方法である。この方法では「思う」と「出来る」が1つの類似性を示している。また「軽い」「速い」「感じる」という言葉の間には太い線で表され、強い関連性があることを示していることがわかる。

7. 考察

RMeCab と KH Coder の結果 (Table. 3) では、「ない・無い」が形態素の頻度として高いが、顧客（消費者）の購入要因やロイヤリティーをつけるための効果的な文章を調査する段階でコンピュータ等、機械での解析では判別が難しい形態素があることが判明した。

また「必要」や「問題」の形態素は、判断の仕方によっては購入要因だけでなく、製品・商品の「デメリット」、「機能不足」等を口コミ、レビューに記載することで、AISAS モデルの「行動段階」のプロセスである「Search（検索）」のプロセスで製品に関する情報や口コミ・レビューの評価内容を見て、購入後に「Share（情報共有）」をすることで、その製品の価値やロイヤリティーの向上に役立てられ、Web ショッピングの売上向上に繋がることが考えられた。「満足」や「必要」等の形態素と製品に関して評価が記された文章との関連性が、顧客の購入要因に繋がると考えられる。

8. おわりに

問題点として、機械では判別できない「形態素」をどのように購入要因として引き出すかについてだが、現状は人間が

文章の前後を見て判断するに他ならない。人間が大量のテキストデータを形態素ごとに解析するのは時間が掛かることから、コンピュータ等の機械を利用してある程度まで分解させ、絞込みを行い、各形態素で考察するのが一般的と言える。

今回の解析結果では1つの製品・商品に対してのレビュー情報が少なく「Mac Book air」全体として調査を行った為、レビューに対する調査結果までであるが、他社製品・商品の「比較」がその製品ごとに関して使われる文章の手法に違いがあるのではないかと考えられる。また口コミやレビューの「信頼性」や「星の数」と「5段階評価」等、形態素と評価による数値との関連性や相関性を見ていきたい。更に情報量を増やすために、PHP（スクリプト言語）とAPIを利用したプログラムを作成し、より多くのデータ収集ができるようにしていきたい。

参考文献

- 1) 上田太一郎, 事例で学ぶ テキストマイニング, 共立出版, 2011年, p. 1
- 2) なぜテキストマイニングは必要か?
<http://www.msi.co.jp/tmstudio/whyTextMining.html>
- 3) 高価な物でも売れる事実と安い物でも売れない事実
<http://www.medi-graph.com/contents/essay/manual/004.html>
- 4) 石田 基広, Rによるテキストマイニング入門, 森北出版, 2008, 12. 22, p45
- 5) KH Coder Index Page
<http://khc.sourceforge.net/index.html>
- 6) クラスター分析とは | 市場調査・アンケート調査のマクロミル
<http://www.macromill.com/landing/words/b003.html>
- 7) Rと多次元尺度法
<http://mj.in.doshisha.ac.jp/R/27/27.html>
- 8) 松本, 製品の顧客評価によるWebショッピングの売上向上戦略, 他 第17回日本情報ディレトリ学会全国大会研究報告予稿集, 2013年 9. 15 p. 63