

広告手法の特徴に基づくクチコミの分類に関する研究

北野光一[†] 寺口敏生[†] 田中成典[‡] 今城彰子[†] 柳田尚明[‡]

関西大学大学院総合情報学研究科[†] 関西大学総合情報学部[‡]

1. はじめに

近年, CGM (Consumer Generated Media) に投稿されるクチコミは, 消費者の購買意思決定に強い影響[1]を及ぼすため, 新しい広告媒体として企業の注目を集めている. 企業が広告を行う場合には, 消費者が商品を認知してから購買に至るまでの過程を複数の段階に分類した購買意思決定過程モデル[2]に基づき, 段階に応じて適切な特性を持つ広告種別を提示し, 消費者の購買意欲を上げることが重要[3]である. そのため, 企業がクチコミを販売促進に活用する場合にも, 購買意思決定過程モデルを考慮してクチコミを分類し, 消費者の段階に応じた特性を持つクチコミを消費者に提示する必要がある. クチコミを分類する既存研究[4][5]では, クチコミを好評文と不評文の2種類に分類している. しかし, 既存研究は, 購買意思決定過程モデルを考慮しておらず, 広告に活用することを目的としたクチコミの分類に適していない. そこで, 本研究では, 企業によるクチコミの活用の支援を目的に, 購買意思決定過程モデルの各段階に対して購買意欲向上に有効な広告種別にクチコミを分類する手法を提案する.

2. 研究の概要

本研究では, 購買意思決定過程モデルの各段階に対して購買意欲向上に有効な広告種別にクチコミを分類する手法を提案する. 本研究で用いる広告種別毎の特徴を表1に示す. 本研究では, 広告の特徴として, 網羅度と詳細度という2つの指標を導入する. 網羅度は, 内容の広さを示す指標であり, 詳細度は, 内容の深さを示す指標である. 本システムの概要を図1に示す. 本システムは, 1) 学習機能, 2) 分類機能により構成される. 入力データは, 対象商品のクチ

表1 広告種別毎の特徴

購買意思決定過程モデルの段階	有効な広告	広告の特徴
未認知	TVCM	網羅度: 低 詳細度: 低
認知・関心	カタログ	網羅度: 高 詳細度: 高
確信	チラシ	網羅度: 高 詳細度: 低

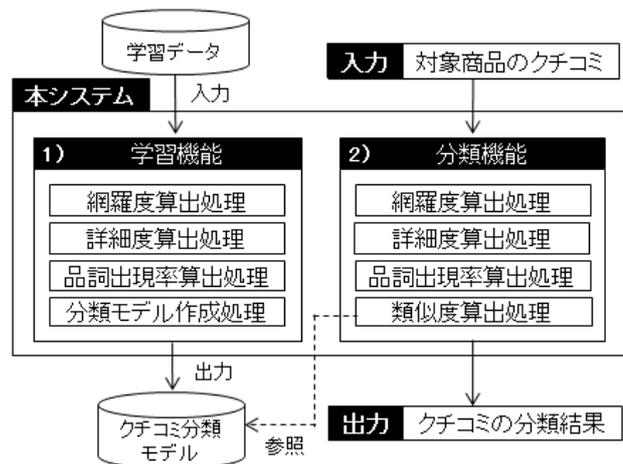


図1 本システムの概要

コミとし, 出力データは, クチコミの分類結果とする.

2.1 学習機能

本機能では, クチコミ毎の網羅度, 詳細度と品詞出現率を算出し, クチコミ分類モデルを構築する. 学習データは, 表1の広告種別に手作業で分類したクチコミを用いる. まず, 網羅度算出処理では, クチコミ中に含まれる商品の性能に関する単語の割合を網羅度として算出する. 次に, 詳細度算出処理では, クチコミ中に含まれる商品の性能に関する単語の関連語の割合をクチコミの詳細度として算出する. そして, 品詞出現率算出処理では, 各クチコミの品詞抽出特徴を抽出するために, 各品詞の品詞出現率を算出する. 利用する品詞は, 広告種別毎の品詞出現率を分析し, 出現率が大きく異なっていた品詞の上位5件を対象とする. 最後に, 分類モ

Research on Classification of Word of Mouth Based on Advertising Characteristic

[†] Koichi Kitano, Toshio Teraguchi, Akiko Imagi, Graduate School of Informatics, Kansai University, 2-1-1 Ryouzenji-cho, Takatsuki-shi, Osaka 569-1095, Japan

[‡] Shigenori Tanaka, Naoaki Yanagida, Faculty of Informatics, Kansai University, 2-1-1 Ryouzenji-cho, Takatsuki-shi, Osaka 569-1095, Japan

デル作成処理では、これらの解析結果を基に広告種別毎の特徴ベクトルを作成し、クチコミ分類モデルを構築する。

2. 2 分類機能

本機能では、学習機能にて構築したクチコミ分類モデルを用いて、クチコミを類似した広告種別に分類する。まず、入力したクチコミの網羅度、詳細度と品詞出現率を算出し、特徴ベクトルを作成する。次に、類似度算出処理では、広告種別毎の特徴ベクトルとクチコミの特徴ベクトルとのコサイン距離をそれぞれ算出し、最も類似した広告種別にクチコミを分類する。

3. システムの実証実験と考察

本システムの有用性を実証するために、価格.comの「クチコミ掲示板」を解析対象として、本提案手法にて用いた網羅度と詳細度の評価指標としての妥当性とクチコミの分類精度について実験を行う。

3. 1 実証実験

本実験では、価格.comのカテゴリの内、ノートPCに関するクチコミを収集し、手作業で表1の広告種別に分類したクチコミを各300件ずつ合計900件用意する。まず、網羅度と詳細度の評価指標としての妥当性に関する実験では、クチコミの分類精度が最も高くなる場合の広告種別毎の網羅度と詳細度の値を最適値として算出する。そして、表1の広告種別の特徴と、網羅度と詳細度の最適値を比較し、一致するかどうかを評価する。分類に用いる指標は、網羅度と詳細度の2つとし、入力データは、用意したクチコミ900件を用いる。次に、クチコミの分類精度に関する実験では、本システムにより出力した分類結果の適合率、再現率とF値を算出することによりクチコミの分類精度を評価する。本実験で分類に用いる指標は、網羅度、詳細度と5種類の品詞の品詞出現率の合計7つとし、用意したクチコミの内、840件を学習データ、60件を実験データとして用いる。

3. 2 結果と考察

網羅度と詳細度の最適値を表2に示す。表1の広告種別の特徴と表2の最適値を比較すると、概ね傾向が一致していることから、本提案手法で用いた網羅度と詳細度が広告種別毎の特徴を表す指標として有用であることが分かった。次に、クチコミの分類結果を表3に示す。全体の分類精度では、F値が0.79であることから、本提案手法で用いた広告種別毎の特徴がクチコミを分類する場合に有効であることが分かった。また、広告種別毎の分類結果を目視で確認した

表2 網羅度と詳細度の最適値

	網羅度	詳細度
TVCM	0.015	0.018
チラシ	0.045	0.000
カタログ	0.047	0.032

表3 クチコミの分類結果

	適合率	再現率	F値
TVCM	0.94	0.68	0.79
チラシ	0.84	0.73	0.78
カタログ	0.68	0.95	0.78
全体	0.79	0.79	0.79

ところ、カタログがTVCMに誤分類される場合が多いことが分かった。これは、網羅度、詳細度と品詞出現率では、カタログとTVCMの特徴差を捉えきれていないことが原因だと考えられる。この問題については、広告とクチコミの特徴を分析し、広告とクチコミに共通して表れる特徴を用いることで解決できると考えられる。

4. おわりに

本研究では、購買意思決定過程モデルの各段階に対して購買意欲向上に有効な広告種別にクチコミを分類する手法の提案を行った。そして、実証実験の結果、網羅度と詳細度が広告種別毎の特徴を表す指標であり、クチコミを分類する際に有用であることから、本システムの有効性を実証した。今後は、広告とクチコミに表れる特徴を分析し、本提案手法で用いた指標に加えて、広告とクチコミに共通して表れる特徴を表す指標を検討することで、分類精度の向上を目指す。

参考文献

- [1] 青木幸弘, 上田隆穂: マーケティングを学ぶ<下>売れ続ける仕組み, 中央経済社, pp.186-210, 2009.12.
- [2] Engel, J., Blackwell, R. and Miniard, P.: Consumer Behavior, South-Western College Publishing, Division of Thomson Learning, 2000.8.
- [3] Schullz, D., Tannenbaum, S. and Lauterborn, R.: Integrated Marketing Communications, McGraw-Hill, 1993.1.
- [4] 安村禎明, 坂野大作, 上原邦昭: 評判情報のレベルを考慮した評価文書の分類と評価情報の信頼性評価への応用, 自然言語処理研究会研究報告, 言語処理学会, Vol.14, No.3, pp.297-313, 2007.4.
- [5] 池田大介, 鷹村大也, 奥村学: 単語極性判定モデルによる評価文分類, 人工知能学会論文誌, 人工知能学会, Vol.25, No.1, pp.50-57, 2010.7.