

令和5年
修士論文草稿

タイトル

指導教員：相馬 隆郎

東京都立大学大学院
電子情報システム工学域

学修番号：22861651

氏名：西原涼介

論文要旨

ここから論文要旨

目 次

第I部

はじめに

1 研究背景

近年, Amazon や楽天市場などの大手 EC サイトをはじめ、数多くの EC サイトが普及し、その利用者も急増している。そして、商品を購入する際に EC サイトのレビューを参考に行っている利用者の割合は約 70 % と言われていて、その中でもレビューの信頼性を重要視している人が多いことが明らかになっている。また、多くの企業にとって、EC サイトのレビューからユーザーの嗜好や意見を分析し、マーケティングに活用することが重要な課題となっている。そのため、EC サイトのレビューの信頼性や参考になるかどうかを評価する評判分析や口コミ分析、レビューを様々なトピックに分類する文書分類に関する研究が多く行われている。例えば、関連研究の項で詳しく紹介する「機械学習を用いた自然言語処理による商品レビューの評価」[1] では、Amazon の商品レビューを機械学習を用いて参考になる順に並びかえるシステムの構築、及びその評価に関する研究を行っている。また近年では、従来の EC サイトや商品の Web ページ以外にも、YouTube のような動画投稿サイトや X(旧 Twitter) や Instagram などの SNS で自社製品・サービスの宣伝を行う企業が増えてきている。それにつれて、商品を購入する際に SNS や YouTube 上でその商品を宣伝している投稿を参考に行っている人も増加している。そのため、SNS や YouTube 上の広告に対するユーザーのコメントも、他のユーザーが商品の購入を検討する際の重要な判断材料になり得ると考えられる。つまり、SNS や YouTube 上での商品の宣伝に対するコメントは、EC サイトのレビューと同等の機能を持ち、その信頼性や参考になるかどうかが必要になるため、評判分析や文書分類の研究の対象になると考えられる。ここで、SNS や YouTube は商品レビューのページとは異なり、誰でも気軽にコメントを投稿できたり、その投稿内容も自由という特性上、商品やサービスに関係ないコメントが多数存在する。

そこで、本研究では分析対象を YouTube 上で自社製品やサービスを宣伝している動画に対するユーザーのコメントとし、トピックモデルの一種である Biterm Topic Model による商品に関するトピック抽出を用いて、その動画に対するユーザーのコメントから、宣伝している商品やサービスに対して関連性が高いコメントを抽出するシステムの作成、及び作成したシステムの人手に対する精度の検証を行った。

本論文の第 I 部では、EC サイトのレビューにおける評判分析やトピックモデルを用いた文書分類に関する関連研究の紹介、また本研究の研究目的を明確に説明する。第 II 部では、本研究で用いる二つのトピックモデルの説明、及び提案手法のシステムや実装方法について説明する。第 III 部では、実際の YouTube 上の動画に対するコメントを用いた実験結果を述べる。第 IV 部では、実験結果をもとに考察した提案手法の有効性や将来性について述べる。

2 関連研究

本研究を進めるにあたり、研究テーマの方向性決めや研究課題の発見、及び本研究で用いている技術に関して参考にした論文を4つ紹介する。

2.1 機械学習を用いた自然言語処理による商品レビューの評価 [1]

この研究では、ユーザーが商品レビューを読んで参考になったかどうかを評価する機能が備わっていないECサイトの場合に、数多くあるレビューから参考になる情報を探す必要がある問題に着目し、機械学習を用いた自然言語処理の手法で分析、評価を行い、レビューを参考になる順番に並び替えるシステムの構築を目的としている。そして並び替えた順番が正しいかどうかを評価するために、クイックソートを利用した新しい評価法であるQE法を提案している。この研究のシステム図が図1である。

2.2 関連研究2

関連研究2の始まり

3 研究目的

研究目的

第II部

提案手法

4 トピックモデル

トピックモデルの説明

4.1 Latent Dirichlet Allocation

LDA の説明

4.2 Biterm Topic Model

BTM の説明

5 提案システム

提案システムの概要、図

5.1 データ収集

5.2 前処理手法

第 III 部

実験結果

6 実験目的、仮説