

サービス評価指標予測による重要意見抽出効果の検証

矢口 一晟 櫻井義尚
明治大学 総合数理学部

1. はじめに

近年普及している SNS にはユーザーから発信される様々な投稿データが蓄積されており、その中には企業の商品・サービス改善などに有益な情報が存在し、マーケティング調査に活用されている。しかしながら、そのような情報は膨大な投稿の中の一部であり、人手でそれを探し出すのは大変な人的コストを要する。そこで本研究では、機械学習を用いてフィルタリングする事で有効な意見(重要意見)を含む投稿の抽出を効率化する手法を提案し、その有効性を検証する。具体的には、まずツイートデータから意見投稿を分類する意見抽出モデルにより、意見ツイートを抽出する。次にサービス評価指標予測器を用いて意見の有効度を推定する。本研究では、サービス評価指標予測器の精度検証をした上で、これを用いた有効意見(重要意見)の抽出効果を検証する。

2. 先行研究

SNS には企業のサービスに対してユーザーから発信される膨大な投稿データが蓄積されており、その中には企業のマーケティング調査に有益な情報が存在する。そのため、SNS における意見抽出の研究が進められている[1]。

また、これまでの自然言語処理分野において批評文等のネガポジ判定の問題が扱われてきたが、意見の中でも「要望」「現状認識」「不満」などの意見の役割を担う各部分の抽出はあまり扱われてこなかった。そこで、山本ら[2]は人々が何を求めているかを知るという点から、機械学習モデルのSVM(Support Vector Machines)を用いて、アンケートの自由回答意見の中でも要望が述べられている要望文を抽出する研究を行った。

山本ら[2]の研究では、「要望」に着目し、横浜環状北西線での建設に関する業務活動における要望分抽出を行うシステムを提案したが、本研究では「サービス表現」に着目し意見データの中から企業がマーケティングに必要なサービス向上・改善に向けた有効な意見(重要意見)の抽出を目指した。

3. 提案手法

3.1 実験の流れ

本研究の目的は意見抽出とサービス評価指標の併用から、マーケティング調査に必要なサービス向上・改善に繋がる有効な意見を抽出することである。そこで2段階の実験を行う。

実験1では図1で示す通り、本研究のフィルタの役割を担うサービス評価指標予測モデルの精度検証、実験2では意見抽出とサービス評価指標予測の併用から有効意見抽出効果の検証を行う。

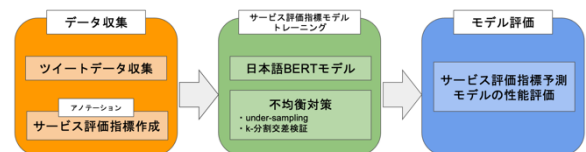


図1 サービス評価モデルの精度検証(実験1)

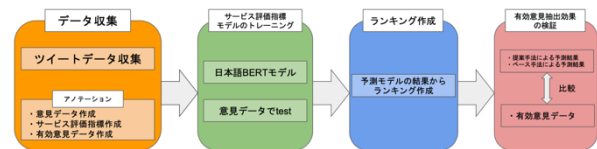


図2 有効意見抽出効果の検証(実験2)

3.2 データ

3.2.1 データセット

本研究では Twitter からの意見抽出対象を大規模サービスサービス施設とし、キーワードをディズニーや USJ, ユニバなどの計 19 施設の名称またはその略称として計 9994 件のツイートを抽出した。

また、ディズニーをキーワードとしたツイートが 3715/9994 件と最多となったため本研究ではディズニーに関するサービス向上・改善に繋がる有効意見抽出を行う。

3.2.2 意見ラベルとサービス評価ラベル

意見抽出で必要となる意見ラベル、サービス評価指標予測モデル構築で必要となるサービス評価ラベ

Extraction of valid opinions that lead to service improvement and enhancement using service evaluation indicators
ISSEI YAGUCHI, YOSHITAKA SAKURAI
School of Interdisciplinary Mathematical Sciences, Meiji University

ルについて述べる。

意見ラベルは抽出したツイートデータが意見であるかどうかを判定する際に用いる指標であり、野崎ら[3]を参考に「要望」「感情・要望」と判定されたものを意見と定義しデータセットの構築を行った。

サービス評価ラベルはフィルタの役割であるサービス評価指標モデル構築に用いる指標でありSERVQUAL[4]という5つの指標をもとに作成した。

3. 2.3 有効意見ラベル

有効意見ラベルは有効意見抽出効果の検証用データに用いられ、有効意見の定義は意見ラベルとは異なり「ディズニーコンテンツを原因としたサービス向上・改善に向けた批評や要望」とした。これは意見と判定されたディズニーのツイート 893 件に対してアノテーションするものであり、検証用データとして利用するため正当性が必要となるため複数でのアノテーションを行った。

3. 3 BERT を用いた分類モデル構築

本研究ではサービス評価指標予測モデル構築において自然言語処理の事前学習モデルBERT(Bidirectional Encoder Representations From Transformers)を用いた。また、多言語モデルであるBERT の特徴として日本語の扱いが難しいという側面から東北大学・乾研究室が公開している日本語 Wikipedia で Pre-Training された日本語版 BERT モデルを使用した。

4. 実験

4.1 実験 1: サービス評価指標予測モデルの精度検証

実験2での意見抽出とサービス評価指標予測による有効意見抽出の検証に先駆けて、日本語版BERTモデルを用いて構築したサービス評価指標予測の有効性を確認するために精度検証を行う。そこで、学習に用いる各サービス評価指標のデータセットは基本的に不均衡データであるため不均衡対策として under-sampling を用いた時とそうでない時で精度比較を行い、検証方法として k-分割検証を用いる。(本研究では k=5) また、分類モデルの評価指標を F1score を用いてモデル評価を行う。

4.2 実験 2: 意見抽出とサービス評価指標予測による有効意見抽出

本研究の目的である意見抽出とサービス評価指標予測の併用からマーケティング調査におけるサービス向上・改善に繋がる有効な意見を抽出できるかを確認する。検証方法としては意見判定されたディズニーのツイート 893 件に有効意見ラベルを用いてアノテーションを行うことで有効

意見を設定し、3.2 で述べた意見ラベルによって抽出された意見をサービス評価指標予測にかけ、最終出力でどの程度網羅できたかを確認する。

ここで、「提案手法の性能確認のために、サービス向上・改善に繋がる意見には投稿者の熱がある」という仮定から感情分析によるベース手法を作成し提案手法との性能比較を行う。

また、性能比較方法として提案手法、ベース手法共に最終出力からランキングを作成し、ランキングの上位検索が有効意見をどの程度網羅できているかで比較する。

5. おわりに

本研究ではマーケティング調査における有効意見の抽出を課題として、意見表現にサービス表現を含んでいる文章が有効な意見であると仮定し、意見抽出とサービス評価指標予測の併用から有効な意見を抽出するシステムを提案した。

今後の課題としてサービス評価指標予測モデルの精度向上に向けた適切な BERT モデルの選択や作成、意見抽出器の作成や不均衡データに対する対策を行う必要がある。

特に不均衡データに対する対策に関しては、サービス評価指標予測モデルの精度向上を図るため各サービス表現をより多く含むツイートデータを取り直すことで有効意見抽出効果の有効性を検証する。

参考文献

- [1]川島崇秀, 佐藤哲司, 神門典子: "Twitter からの消費者ニーズの抽出手法に関する提案", 第 8 回データ工学と情報マネジメントに関するフォーラム, (2016)
- [2]山本瑞樹, 乾孝司, 高村大也, 丸元聡子, 大塚裕子, 奥村学: "文章構造を考慮した自由解答意見からの要望抽出", 言語処理学会 12 回年次大会併設ワークショップ「感情 評価 態度と言語」, (2006)
- [3]野崎雄太, 櫻井義尚: "Twitter からの意見抽出モデル構築のための教師データ作成手法", 情報処理学会論文誌数理モデル化と応用 (TOM) pp. 69-83, (2020)
- [4] A Parsu Parasuraman, Valarie A. Zeithaml, Leonard L Berry: "SERVQUAL: A multiple-Item Scale for measuring consumer perceptions of service quality", Journal of Retailing, (1988)