感情分析を用いた口コミの正当性判定手法に関する研究

髙橋, 藍子 / TAKAHASHI, Aiko

(出版者 / Publisher)
法政大学大学院理工学研究科
(雑誌名 / Journal or Publication Title)
法政大学大学院紀要. 理工学研究科編
(巻 / Volume)
65
(開始ページ / Start Page)
1
(終了ページ / End Page)
6
(発行年 / Year)
2024-03-24
(URL)
https://doi.org/10.15002/00030746

感情分析を用いた口コミの正当性判定手法に関する研究

A STUDY ON METHODOLOGY FOR JUDGING THE LEGITIMACY OF WORD-OF-MOUTH USING SENTIMENT ANALYSIS

髙橋藍子
Aiko TAKAHASHI
指導教員　金井敦

法政大学大学院理工学研究科応用情報工学専攻修士課程

In recent years, many people are influenced by posts on social networking services (SNS) when considering the purchase of products and other items, and the influence of word of mouth is growing. Because of this trend, SNS marketing is becoming more important for companies, and there is a growing need to analyze useful information from word-of-mouth that can be utilized for product development and other purposes. However, word-of-mouth includes malicious data such as slanderous remarks, and such data must be classified and excluded. Therefore, sentiment analysis, which can determine whether the content of word-of-mouth is positive or negative, is considered to be effective.

In this study, we propose a method that uses sentiment analysis and machine learning to exclude malicious content from word-of-mouth and extract complaints that lead to points for improvement. We demonstrate the effectiveness of the proposed method by using about 12,000 postings for evaluation and analyzing them with well-defined features.

***Key Words** : Sentiment analysis, machine learning, SNS*

１．はじめに

近年，商品などの購入検討の際に SNS 上の投稿に影響される人が多く，口コミの影響力が大きくなっている．このような傾向から，企業側では SNS マーケティングが重要視されており，口コミから商品開発などに活かせる有益な情報を分析する必要性が高まっている．しかし，口コミの中には誹謗中傷などの悪意のあるものも含まれており，そのようなデータを分類して除外する必要がある．そこで，口コミの内容がポジティブかネガティブかを判別できる感情分析が有効と考えられる．

本研究では，感情分析と機械学習を用いて，口コミの中から悪意のあるものを除外し，改善点に繋がる不満を抽出する手法を提案する．この研究によって，有益な情報を得るとともに，商品やサービスのイメージの低下に繋がる不満を把握し，消費者のニーズに合わせて改善できるようになると考えられる．また，約 12,000 件の SNS 上の投稿を用いて特徴量を明確にして分析し，提案手法の有効性を検証する．

２．研究背景

近年，購入商品や購入店舗に関する口コミを SNS に投稿する人が増加している．アライドアーキテクツ株式会社の行ったアンケートによると，Instagram では 57.0%，Twitter では 53.5%の人が感想を投稿した経験があった [1]．Twitter は 2023 年 7 月に名称を X に変更しているため，本稿では以降，X(旧 Twitter)と表記する．アジャイルメディア・ネットワーク株式会社の行ったアンケートでは，商品などの購入検討の際に影響を受ける口コミとして，SNS 投稿が 50%と最も多く，次いで友人・知人が 43%，家族が 37%であった [2]. SNS 上の口コミの影響力が大きくなっている．

2021 年に行われたアライドアーキテクツ株式会社の調査によると，2020 年から 2021 年にかけて SNS マーケティング施策への予算の変化について，「非常に増加した」と回答した企業が 25.0%，「やや増加した」と回答した企業が 45.4%と，回答した企業の 7 割以上が予算を増やしていた [3]．2022 年の SNS マーケティング予算に関しても，回答した企業の 7 割近くが増加させる予定と答えた．

このような傾向から，企業側は SNS マーケティングを重要視しており，口コミから販売戦略や商品開発に活かせる有益な情報を分析する必要性が高まっている．

しかし，口コミの中には誹謗中傷などの悪意のあるものも含まれており，そのようなデータを分類して除外する必要がある．そこで，口コミ内容がポジティブかネガテ

ィブか判別できる感情分析が有効と考えられる．

感情分析とは，文章に含まれている単語や表現を分析し，それぞれにポジティブかネガティブかのスコア付けをすることで，文章全体がポジティブかネガティブか，どちらでもないニュートラルか判断する手法である．さらに，ポジティブかどうかだけではなく，喜び，悲しみ，怒りなどの感情の状態を識別することもできる．感情分析は，顧客の満足度の分析や，デマの検知，ブランドイメージの把握などに活用されている．

本研究では感情分析と機械学習を用いて，口コミの中から悪意のあるものを除外し，改善点に繋がる不満を抽出する手法を提案する．この研究によって，有益な情報を得るとともに，商品やサービスのイメージの低下に繋がる不満を把握し，消費者のニーズに合わせて改善できるようになると考えられる．

３． 関連研究

感情分析を用いた関連研究として，ハリウッド映画への評価の分析やブランドへのイメージの分析などが挙げられる．

Umesh Rao Hodeghatta は，ハリウッド映画に関する X(旧 Twiiter)の投稿をポジティブな感情，ネガティブな感情，および客観的な情報として分類し，様々な国の様々な地域の人々が表明する感情，意見を分析する手法を提案している [4].

Dua'a Al-Hajjar らは，ブランドへのイメージの分析をする研究において，ブランドに関連する X(旧 Twiiter)の投稿に含まれる感情がポジティブかネガティブか判断するだけではなく，さらに細かい感情に分類する手法を提案している [5]．この手法では，ポジティブな感情から期待，信頼，喜びの感情のスコアを算出し，ネガティブな感情から怒り，恐れ，悲しみ，嫌悪感の感情のスコアを算出している．

４． 研究目的

3 章で述べたように，感情分析を用いた関連研究では，投稿内容を感情ごとに数値化することなどはできている．しかし，その中に悪意や不満といった感情は含まれておらず，悪意と不満を分類することはできていない．

そこで，X(旧 Twiiter)に投稿された口コミをさらに悪意と不満に分類する手法を提案する．本研究では感情分析と機械学習を用いて，口コミの中から悪意のあるものを除外し，改善点に繋がる不満を抽出する手法を提案する．

本研究の目的は，口コミから有益な情報を得ることにより，商品やサービスのイメージの低下に繋がる不満を把握し，消費者のニーズに合わせて改善できるようにすることである．

５． 提案手法

本研究では，感情分析と機械学習を用いて，口コミの中から悪意のあるものを除外し，改善点に繋がる不満を抽出する手法を提案する．提案手法のイメージを以下の図 1 に示す．

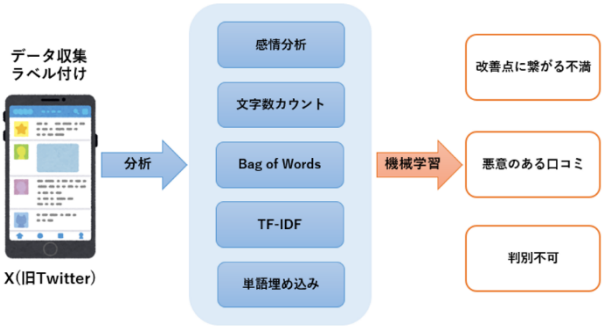

図 1　提案手法のイメージ

はじめに，X(旧 Twitter)からデータを収集し，ラベル付けする．次に，感情分析をはじめとした 5 つの手法でデータを分析し，特徴量を抽出する．その結果をもとに機械学習を用いて，改善点に繋がる不満と悪意のある口コミ，判別不可の 3 つに分類する．

（１）分析データの収集とラベル付け

分析データの収集において X(旧 Twitter)を使用したのは，文字が主体で感情分析を行いやすいからである．他の SNS として，Facebook や Instagram が挙げられるが，画像が中心となる投稿が多いので分析手法には適さないと判断した．分析データとして，スマートフォンの Xperia に関連する投稿を 6,319 件，Google Pixel(以下 Pixel)に関連する投稿を 6,255 件収集し，合計約 12,000 件のデータを収集した．

ラベル付けは手作業によって行い，不満，悪意，判別不可の 3 つに分類した．判別不可には，不満でも悪意でもないその他のデータを分類した．悪意に分類する際には，改善点につながる内容が含まれていないこと，または攻撃的な表現が含まれていることを判断基準とした．ラベル付けの結果を以下の表 1 に示す．

表 1　ラベル付けの結果

	不満	悪意	判別不可	合計
Xperia	937	53	5,329	6,319
Google Pixel	674	74	5,507	6,255
総データ	1,611	127	10,836	12,574

（２）特徴量の抽出

特徴量の抽出には，感情分析，文字数のカウント，Bag of Words，TF-IDF，単語埋め込みの 5 つの手法を用いた．

感情分析では，Google Cloud Platform が提供する Natural Language API を使用して．投稿内容の感情のスコアと感情の強さを算出した．この API では投稿内容全体の感情について-1.0~+1.0 でスコア付けし，感情の強さを 0 以上の数値で算出する．感情のスコアの分布を以下の図

2 に示す.

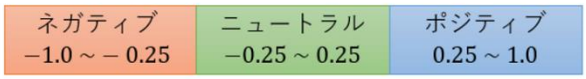

図 2　Natural Language API の感情のスコアの分布

また，不満を述べているものは文章が長くなる傾向があるのに対して，悪意のあるものは端的に述べられている傾向があったため，文字数をカウントした.

Bag of Words, TF-IDF, 単語埋め込みの 3 つの手法では，形態素解析を行い，名詞，動詞，形容詞のみを分析の対象とした．形態素解析には MeCab と Unidic 辞書を用いた [6] [7]．そして，Bag of Words と TF-IDF を用いて，悪意に分類されたデータの中で TF-IDF 値の高い 20 単語の出現頻度と TF-IDF 値を求めた．単語埋め込みでは FastText というモデルを使用して，投稿内容に含まれる単語のベクトルの平均をとり，投稿内容をベクトル化した．さらに，投稿内容のベクトルを主成分分析を用いて 5 次元に次元削減した.

これらの 5 つの手法で抽出した特徴量は以下の通りである.

- 感情のスコア
- 感情の強さ
- 文字数
- 悪意のデータ内で TF-IDF 値の高い 20 単語の出現頻度
- 悪意のデータ内で TF-IDF 値の高い 20 単語の TF-IDF 値
- 投稿内容のベクトル

（３）機械学習による分類

収集したデータを学習に使用する訓練データと，精度の評価に用いるテストデータに分割した．Xperia のデータは 75%を，Pixel のデータと総データは 70%を訓練データとした. 学習はナイーブベイズ, ロジスティック回帰, ランダムフォレストの 3 つの手法で行い，抽出した特徴量を正規化し，最小値が 0，最大値が 1 になるようにして使用した．訓練データで学習した後にテストデータで検証し，それぞれの分類結果を比較して評価した.

６．評価検証

Xperia，Pixel，総データのデータセットに対してそれぞれ提案手法を用いて，不満，悪意，判別不可の 3 つのクラスに分類した.

（１）評価指標

評価指標には分類結果全体の正解率, 適合率, 再現率, F 値と, 各クラスにおける適合率, 再現率, F 値を用いた. これらの評価指標は混同行列から求め, 全体の適合率, 再現率, F 値は各クラスの評価指標の平均をとることで算出した.

混同行列とは，分類結果の予測クラスと正解クラスを軸とした行列であり，予測クラスと正解クラスの組み合わせごとの分類結果がまとめられている.

正解率とは，全体のデータの中で正しく予測できたデータの割合であり，適合率とは，Positive と予測した中で実際に Positive だったデータの割合である [8]．再現率とは，Positive のデータの中で実際に Positive と予測できたデータの割合である [8]．F 値とは，適合率と再現率の調和平均であり，この 2 つの指標のバランスを表す [8].

（２）悪意のデータ内で TF-IDF 値の高い 20 単語

各データセットの悪意のデータ内で TF-IDF 値の高い 20 単語を，以下の表 2 に示す.

表 2　悪意のデータ内で TF-IDF 値の高い 20 単語

データセット	悪意のデータ内で TF-IDF 値の高い 20 単語
Xperia	pixel, 使っ, iphone, 買わ, 買い, 日本, 悲報, アイドル, 使わ, ネトウヨ, 起用, 広告, 全員, ww, pop, 発狂, シリーズ, 嫌い, ない, 機種
Pixel	iphone, pro, レビュー, tensor, 消費, 採用, 財布, gen, 破壊, snapdragon, 失敗, 発熱, 性能, ポンコツ, 劣る, 高温, 想像, se, 絶する, らいん
総データ	iphone, pro, レビュー, tensor, snapdragon, 財布, 失敗, gen, 採用, 破壊, 消費, 性能, 発熱, 絶する, 想像, 劣る, 高温, ポンコツ, se, 使っ

Xperia のデータセットでは他 2 つのデータセットに比べると, 比較的一般的な単語が多く, ネガティブな単語が少なかった．総データの単語には，Pixel と共通するものが多く含まれていた.

（３）Xperia のデータセットの分類結果

Xperia のデータセットをナイーブベイズ，ロジスティック回帰, ランダムフォレストを用いて分類し, 混同行列と各クラスの適合率，再現率，F 値を算出した．また，各手法の正解率と全体の適合率，再現率，F 値を算出した.

次に，表 1 に示したように，判別不可のデータ数が多く, データがクラスごとに偏っているため, アンダーサンプリングを行った. アンダーサンプリングとは, データ数の多いクラスのデータを削減することにより，各クラスのデータ数の偏りを改善する手法である [9]. この手法を用いて判別不可のデータを 5 割に削減した．その結果を以下の表 3～表 6 に示す.

表 3　アンダーサンプリング後の Xperia における各クラスの評価指標(ナイーブベイズ)

ナイーブベイズ	適合率	再現率	F 値
不満	0.3043	0.0311	0.0565
悪意	0.7500	0.2308	0.3529
判別不可	0.8551	0.9896	0.9174

表 4　アンダーサンプリング後の Xperia における
各クラスの評価指標(ロジスティック回帰)

ロジスティック回帰	適合率	再現率	F 値
不満	0.4748	0.5867	0.5249
悪意	1.0000	0.2308	0.3750
判別不可	0.9253	0.8957	0.9103

表 5　アンダーサンプリング後の Xperia における
各クラスの評価指標(ランダムフォレスト)

ランダムフォレスト	適合率	再現率	F 値
不満	0.4891	0.6000	0.5389
悪意	1.0000	0.2308	0.3750
判別不可	0.9277	0.8994	0.9134

表 6　アンダーサンプリング後の各手法の
Xperia における評価指標

	正解率	適合率	再現率	F 値
ナイーブベイズ	0.8468	0.6365	0.4171	0.4423
ロジスティック回帰	0.8462	0.8000	0.5710	0.6034
ランダムフォレスト	0.8513	0.8056	0.5767	0.6091

（４）Pixel のデータセットの分類結果

Pixel のデータセットの分類した際にも，Xperia と同様にアンダーサンプリングを行った．アンダーサンプリング後の各クラスの評価指標と全体の評価指標をそれぞれ以下の表 7～表 10 に示す.

表 7　アンダーサンプリング後の Pixel における
各クラスの評価指標(ナイーブベイズ)

ナイーブベイズ	適合率	再現率	F 値
不満	0.3333	0.0138	0.0264
悪意	0.6000	0.5294	0.5625
判別不可	0.8818	0.9951	0.9351

表 8　アンダーサンプリング後の Pixel における
各クラスの評価指標(ロジスティック回帰)

ロジスティック回帰	適合率	再現率	F 値
不満	0.5278	0.3486	0.4199
悪意	0.6923	0.5294	0.6000
判別不可	0.9163	0.9598	0.9375

表 9　アンダーサンプリング後の Pixel における
各クラスの評価指標(ランダムフォレスト)

ランダムフォレスト	適合率	再現率	F 値
不満	0.5316	0.3853	0.4468
悪意	0.7500	0.5294	0.6207
判別不可	0.9197	0.9562	0.9376

表 10　アンダーサンプリング後の各手法の
Pixel における評価指標

	正解率	適合率	再現率	F 値
ナイーブベイズ	0.8769	0.6050	0.5128	0.5080
ロジスティック回帰	0.8849	0.7121	0.6126	0.6525
ランダムフォレスト	0.8860	0.7338	0.6236	0.6684

（５）総データのデータセットの分類結果

総データのデータセットの分類した際にも，Xperia，Pixel と同様にアンダーサンプリングを行った．アンダーサンプリング後の各クラスの評価指標と全体の評価指標をそれぞれ以下の表 11～表 14 に示す.

表 11　アンダーサンプリング後の総データにおける
各クラスの評価指標(ナイーブベイズ)

ナイーブベイズ	適合率	再現率	F 値
不満	0.3934	0.0493	0.0876
悪意	0.5000	0.2333	0.3182
判別不可	0.8697	0.9877	0.9249

表 12　アンダーサンプリング後の総データにおける
各クラスの評価指標(ロジスティック回帰)

ロジスティック回帰	適合率	再現率	F 値
不満	0.4591	0.5421	0.4972
悪意	0.7000	0.2333	0.3500
判別不可	0.9282	0.9088	0.9184

表 13　アンダーサンプリング後の総データにおける
各クラスの評価指標(ランダムフォレスト)

ランダムフォレスト	適合率	再現率	F 値
不満	0.4609	0.5688	0.5092
悪意	0.6667	0.2000	0.3077
判別不可	0.9311	0.9045	0.9176

表 14　アンダーサンプリング後の各手法の総データにおける評価指標

	正解率	適合率	再現率	F 値
ナイーブベイズ	0.8606	0.5877	0.4234	0.4436
ロジスティック回帰	0.8561	0.6958	0.5614	0.5885
ランダムフォレスト	0.8556	0.6862	0.5578	0.5782

（６）分類結果の評価

Xperia，Pixel，総データのデータセットでアンダーサンプリングを行い，全てにおいて悪意の分類精度はあまり変わらなかったが，不満の分類精度が向上し，全体の分類精度も向上した．Xperia では，ロジスティック回帰とランダムフォレストを用いた際に，最も高い評価指標の値が得られた．Pixel では，ランダムフォレストで最も高い評価指標の値が得られ，総データではロジスティック回帰で最も高い評価指標の値が得られた．

分類精度については，Pixel が最も高く，次いで Xperia，総データの順に精度が高かった．ナイーブベイズ，ロジスティック回帰，ランダムフォレストの比較をすると，ナイーブベイズの精度が最も低く，その他の 2 つの手法には大きな差はなかった．

７．考察

（１）提案手法の有効性

Xperia，Pixel，総データのデータセットにおける不満，悪意，判別不可の 3 つのクラスへの分類について，3(3)～3(6)に示した通り，一定程度の精度を得られた．Pixel の不満と悪意の分類精度が最も高かったのは，各クラスのデータ数の偏りが最も小さかったことと，特徴量に用いた悪意のデータ内で TF-IDF 値の高い単語に，ネガティブな単語が多く含まれていたためと考えられる．悪意を 1 とした際の各データセットにおけるクラスごとのデータ数の比率は，以下の表 15 の通りとなり，Pixel のデータセットが一番データ数の偏りが小さかった．

表 15　悪意を 1 とした際のクラスごとのデータ数の比率

データセット	不満: 悪意: 判別不可
Xperia	18:1:101
Pixel	9:1:74
総データ	13:1:85

Xperia のデータセットで不満と悪意の精度が低かったのは，悪意のデータ内で TF-IDF 値の高い単語に，比較的一般的な単語が多く，ネガティブな単語が少なかったためと考えられる．総データでも精度が低かったのは，総データと Pixel の悪意のデータ内で TF-IDF 値の高い単語に共通するものが多かったうえに，Xperia の単語は比較的一般的な単語が多かったため，抽出した特徴量では Pixel のデータは分類できるが，Xperia のデータは分類できなかったことが原因と考えられる．

したがって，不満や悪意のデータ数を増加させて，各クラスの偏りを小さくすることによって，分類精度を向上させられると考えられる．さらに，悪意のデータ数が増加することにより，悪意のデータ内で TF-IDF 値の高い単語にネガティブな単語が含まれる割合も高くなり，精度の向上に繋がると考えられる．

以上のことから，本研究の提案手法は，口コミの中から悪意のあるものを除外し，改善点に繋がる不満を抽出することに対して有効と考えられる．

（２）今後の課題と展望

アンダーサンプリングによりクラスごとのデータ数の偏りを改善しようとしたが，悪意のデータ数が他のクラスに比べてかなり少なく，不満のデータ数と判別不可のデータ数の偏りは小さくなったが，悪意のデータとの偏りは大きいままだったため，不満の分類精度のみが向上し，悪意のデータについては変化があまりなかったと考えられる．

データ数の偏りを改善するためには，より多くのデータを収集し，不満と悪意のデータ数を増やすとともに，判別不可のデータ数を減らすことも必要であると考えられる．本研究では，アンダーサンプリングにより判別不可のデータ数を削減したが，より分類精度を向上させるためには，感情のスコアが 1.0 に近い値となる極めてポジティブな内容のデータなどの，不満や悪意との関連性の低いデータを削減することが有効と考えられる．

また，ナイーブベイズの分類精度が他の 2 つの手法に比べて低かったのは，ナイーブベイズではデータの各特徴量に相関がなく独立していると仮定して学習を行うが，本研究では特徴量に用いた単語の共起語を考慮できておらず，各特徴量が独立ではなかったことが原因と考えられる．共起語とは，ある文書の中で特定の単語と同時に使用されることの多い単語のことである．本研究では，特定の商品に関連する投稿を分類対象としたため，より共起語の影響が大きく，特徴量の相関が強かったと考えられる．したがって，本研究の手法には適していなかったと考えられる．

今後は，データ数を増やして精度を向上させるとともに，不満のデータから改善の方向性ごとに分類して，商品開発などに有益な情報を分析する必要があると考えられる．データ数が増加すると，手作業でのラベル付けに膨大な時間がかかるため，ラベル付けの効率化も必要になると考えられる．そのためには，不満や悪意のデータ内でのみ出現頻度の高い単語をそれぞれリスト化し，リスト内の単語が含まれるデータを自動で抽出してラベル付けするようなシステムが有効と考えられる．このシステムを用いると，一部のデータを手作業でラベル付けした後は自動で不満や悪意のデータが抽出されるため，手作業でラベル付けするデータ数を削減することが可能になると考えられる．

８．結論

本研究では，感情分析と機械学習を用いて口コミの中から悪意のあるものを除外し，改善点に繋がる不満を抽出する手法を提案した．提案手法を用いて，X(旧 Twiiter)から収集したデータを不満，悪意，判別不可の 3 つのクラスに分類し，評価検証を通して有効性を確認できた．

今後はデータ数を増やして精度を向上させるとともに，不満のデータから改善の方向性ごとに分類していく．また，本研究ではラベル付けを手作業で行ったが，データ数が増加すると時間がかかるため，ラベル付けの効率化についても取り組む．

参考文献

[1] SMMLab, “【2020 年最新版】5 大 SNS ユーザーによる「SNS をきっかけとした購買行動・口コミ行動調査結果」公開！（Twitter、Instagram、Facebook、LINE、YouTube）,” 2021. [オンライン]. Available: https://smmlab.jp/article/sns-research-2020/. [アクセス日: 8 12 2022].

[2] アジャイルメディア・ネットワーク株式会社, “[AMN 調査リリース] SNS のクチコミが 生活者の購入・来店に与える影響を調査,” 2022. [オンライン]. Available: https://agilemedia.jp/pr/release220926.html. [アクセス日: 5 12 2022].

[3] echoes, “【2021 年版・業種別】企業の「SNS マーケティング実態調査」結果発表！2022 年の見込みも公開～食品・飲料・化粧品・外食・小売～,” 2021. [オンライン]. Available:https://service.aainc.co.jp/product/echoes/voices/0047. [アクセス日: 8 12 2022].

[4] Umesh Rao Hodeghatta, “Sentiment Analysis of Hollywood Movies on Twitter,” 2013.

[5] Dua'a Al-Hajjar , Afraz Z. Syed, “Applying Sentiment and Emotion Analysis on Brand Tweets for Digital Marketing,” 2015.

[6] “MeCab: Yet Another Part-of-Speech and Morphological Analyzer,” [オンライン]. Available: https://taku910.github.io/mecab/. [アクセス日: 3 2 2024].

[7] 言語資源開発センター, “「UniDic」国語研短単位自動解析用辞書 - 言語資源開発センター,” [オンライン]. Available: https://clrd.ninjal.ac.jp/unidic/. [アクセス日: 3 2 2024].

[8] 株式会社セールスアナリティクス, “第 312 話｜統計的機械学習でよく使用される混同行列（Confusion Matrix）と評価指標,” 2022. [オンライン]. Available: https://www.salesanalytics.co.jp/column/no00312/. [アクセス日: 16 2 2024].

[9] 株式会社セールスアナリティクス, “第 364 話｜機械学習の課題: 不均衡データへのアンダーサンプリング解決策,” 2023. [オンライン]. Available: https://www.salesanalytics.co.jp/column/no00364/. [アクセス日: 11 2 2024].