

レーベンシュタイン距離行列及びセンチメント特徴量を用いた スパム判定モデル

赤澤優佑[†] 大槻明[‡]

日本大学経済学部[‡]

1 はじめに

近年、様々な EC サイトの登場により、商品の口コミであるレビューがユーザの購買意欲に大きく関わるようになってきている。しかし、Amazon が 2020 年に 2 億件を超えるスパムが疑われるレビューを削除したと発表したように、スパムレビューを用いてユーザの購買意欲を不正に操作するケースが問題になってきている。ゆえに本研究では、独自にスパムレビューの定義を定め、その定義に当てはまるスパムレビューを検知するためのスクリプトを研究開発した。

2 本研究におけるスパムの定義

本研究では、サクラチェッカーのトップページで紹介されていた、タブレット、電動歯ブラシ、カーチャージャー、電源タップ、スピーカー、体重計の 6 商品を題材としてレビュー分析を行うこととし、係る 2116 件の商品レビューをアマゾンから収集した。そして 80 件のサンプルチェックを通じて、下記①②に示すスパムレビューの特徴を明らかにした。

- ① レビューテキストの一部の文章が類似
- ② 極端に商品を褒めているもしくは貶している内容が含まれている

なお、取得したレビューは、まずサクラチェッカーで商品検索を行い、レーティング情報がある商品レビューを Amazon から取得する形にしているため、サクラチェッカーのレーティングと照合することが可能である。そして、上記①②の特徴を持つレビューのサクラ度は、6 商品の平均で 96%であったため、本節のスパム定義にはある程度の信ぴょう性があると考えられる。

3 提案アプローチ

3.1 はじめに

まず初めに提案アプローチの全体像について述べる。本提案モデルの最初の手順は TF-IDF 分析 (3.2 節) により対象文書の絞り込みを行う、これはスパム判定を行う対象レビューの絞り込むことが目的である。次に、レーベンシュタイン距離を用いた分析 (3.3 節) を行うことで、前章の①の判定を行う。そして最後に、センチメント分析 (3.4 節) を行うことで前章の②の判定を行い①②の結果から総合的にスパム判定を行う。

3.2 TF-IDF 及びコサイン類似度を用いた類似文書抽出

まず、レビューテキストをデータフレームに結合するが、この時に 20 文字未満のレビューを除外する。これは、短すぎるレビューは実際の類似度に関わらず類似レビューとして抽出されてしまう (例: 「この商品良い!」と「この商品良いです」) からであり、前章の 2116 件のレビューデータをサンプルチェックした結果を受けて、20 文字という基準を設定した。その後レビューデータのベクトル化のために、形態素解析を行った後で文章ごとに TF-IDF 分析を行いそのスコアを形態素に付与することでベクトル化する。ベクトル化されたレビューのイメージを表 1 に示す。このように結果のベクトルは文章ごとに各単語の TF-IDF 値となる。

表 1. TF-IDF でベクトル化されたレビューイメージ

| レビュー | 良い | 買い物 | が | できました |
|--------|------|------|------|-------|
| TF-IDF | 3.78 | 3.70 | 4.70 | 5.67 |

そして、この文章ベクトルを対象にコサイン類似度 [1] を用いて 2 レビュー間の類似度を求める。本研究ではこの 2 レビュー間のベクトルの類似度が高いレビュー上位 50 ペア、計 100 件を抽出して次節の分析を行った。

3.3 レーベンシュタイン距離を用いたテキストの一部が類似する 2 レビューの抽出

前節で抽出されたレビューに対し、レーベンシュタイン距離を用いてレビューテキストの一部が類似する 2 つのレビューの抽出を行う。これは前章の①の工程である。具体的には、下記の通りレビューを Statement 単位で分割したうえで Statement 単位でレーベンシュタイン距離を求める。なお分割基準には「,」「.」「、」「。」「!」「!」を用いた。

【Statement 分割前】

7 インチから 10 インチに買い換えました。デフォルトでの表示がなかったのですが WiFi につないでいるとアップデート画面が出てアップデートすると日本語の選択項目が出てきて日本語化できました。重さはずっとくる感じです。充電は手持ちの PD 充電器が使えませんでしたので付属の物を使用しています。

↓

【Statement 分割後】

- ・ 7 インチから 10 インチに買い換えました。
- ・ デフォルトでの表示がなかったのですが WiFi につないでいるとアップデート画面が出てアップデートすると日本語の選択項目が出てきて日本語化できました。

- ・ 重さはずしっとくる感じです。
- ・ 充電は手持ちの PD 充電器が使えませんでしたので付属の物を使用しています。

工程としては、前節で抽出したレビューデータの 1 件に対して他の 99 件のレビューとのレーベンシュタイン距離の算出を行い、その工程をすべてのレビューで行う。そしてレーベンシュタインスコアが低い順に表示する。

3.4 センチメント分析による極端にポジティブ/ネガティブなレビューの抽出

前節で抽出されたレビューテキストの一部の文章が類似する 2 レビューに対し、極端に褒める、または貶すレビューかどうかを判定するために極性分析を行う。これは第 2 章で定めたスパム定義の②に該当する。極性分析に使用する辞書には単語感情極性対応表[2]を使用した。

3.5 スпам判定

本提案スクリプトの最終的な出力イメージは表 2 に示す通りである。表 2 は、提案スクリプトが出力した 1 番目の（つまりレーベンシュタインスコア=0 の）2 レビュー

表 2. 本スクリプトの出カイメージ

| | Levenshtein distance Score | Review | Sentiment Score |
|-------|----------------------------|--|--|
| レビュー1 | 0.0 | バッテリーも長持ちして、5-6 時間は持ちます。これだけ持ったら十分ですね。部屋を歩き来する時も持っていきます。5000 円の割引も嬉しかったです。 | ['長持ち', '十分', '部屋', '行き来', '割引', '嬉しい'] [0.967016 -0.313183 -0.947083 -0.578158 -0.388726 0.998871] Total Score : -0.261263 |
| レビュー2 | | バッテリーも長持ちして、5-6 時間は持ちます。これだけ持ったら十分ですね。部屋を歩き来する時も持っていきます。5000 円の割引も嬉しかったです。 | ['長持ち', '十分', '部屋', '行き来', '割引', '嬉しい'] [0.967016 -0.313183 -0.947083 -0.578158 -0.388726 0.998871] Total Score : -0.261263 |

4 考察

レーベンシュタインスコアの上位 50 件を調査したところ、極端に褒めているスパムレビューを 1 件、極端に貶しているスパムレビューを 5 件をそれぞれ検出することができた。極端に褒めているスパムレビューは下記に示す通り問題無く検出されていたが、

| Distance score | Sentiment score | Review |
|----------------|-----------------|----------------------------|
| 0.00 | 0.74 | タッチの反応が良く、また画像もきれいなので満足です。 |
| | 0.74 | 上と全く同じため割愛 |

しかし、極端に貶しているスパムレビューについては、5 件中 3 件は下記に例示するように商品を褒めているレビューであった。この原因の一つには極性分析に使用した辞書の問題が考えられるため、今後の課題として検討していきたい。

| Distance | Sentiment | Review |
|----------|-----------|--------|
| | | |

であるため内容が全く同じであるが、スパム判定は「Levenshtein distance Score」列のスコアと「Sentiment Score」列のスコアを総合的に判断して行う。具体的には、レーベンシュタインスコアの上位 50 のレビューペアにおける①Levenshtein distance Score 及び②正の Sentiment Score (Total Score) と ③負の Sentiment Score (Total Score) の平均値はそれぞれ①0.54, ②0.43, ③-2.84 であったため、本研究ではこの①②の平均値を超えるペアレビューを極端に褒めているスパムレビュー、①③の平均値を超えるペアレビューを極端に貶しているスパムレビューと判定することとした。なお、この閾値は本研究で実証したデータ用に算出したものであり、また、レーベンシュタインスコアの上位 50 とした理由も、51 位以降になると、センチメントスコアの平均値が極端に低くなくなったためである。つまり、このスパム判定に関する最適な閾値については、他の商品もいくつか実証したうえで引き続き検討する必要があると考える。

| score | score | |
|-------|--------|--|
| 0.00 | -10.66 | 登録もさくさく進められたし乗ればオンになるしとても快適に始められました。簡単な説明書が入っていますがそこからコード読み取ってアプリを DL して書いてあるとおりにすれば誰でもすぐ設定とリンクできると思います。病院で測るほどの正確性はないにしても確実に目安にはなると思うので色々考えていこうと思います。乗っただけで全部アプリが管理してくれるのが便利ですね |
| | -10.66 | 上と全く同じため割愛 |

参考文献

- [1] 難波英嗣：テキスト間の類似度の測定，情報の科学と技術，70 巻，7 号，pp.373～375，2020.
- [2] 高村大也，乾孝司，奥村学 “スピンモデルによる単語の感情極性抽出”，情報処理学会論文誌ジャーナル，Vol.47 No.02 pp. 627--637，2006.