

SNS のコメントにもとづいたユーザ感情分析

趙 雨陽[†] 上野翔慶[†] 大塚 真吾^{†‡}

^{† ‡} 神奈川工科大学情報学部 〒243-0292 神奈川県厚木市下荻野 1030

E-mail: [†] {s2085001, s1821004}@cco.kanagawa-it.ac.jp, ^{† ‡} otsuka@ic.kanagawa-it.ac.jp

あらまし SNS 上では人々が日々の出来事や個人的な意見・感情などの情報を発信しているため、感情に関するデータの収集が容易である。感情が含まれたデータを分析することで世間の声を把握することができ、自社製品の改善、戦略立案、マーケティング活動など幅広い分野に活用されている。本研究ではソーシャル・ネットワーク・サービスにおけるユーザのレビューコメント等を収集し、大量なコメントの中の感情について、効率的に分析する手法の提案を行う。

キーワード SNS 感情分析 機械学習

1. はじめに

現在、個人の生活や会社において Web を利用する機会が増えている。Web 上において、ユーザが商品やアプリケーションを購入する際に、他の購入者のレビューコメントを参考にすることが多く、購入の判断に大きな影響を与えている^[1]。さらに、ある事件に対するソーシャル・ネットワーク・サービス (SNS) でのユーザのコメントは、世論の傾向にも影響を与える可能性がある^[2]。影響を与えるコメントの多くは感情を含んでいることが多いため、これらの情報を効率的に分析することが重要である。また、中国語と日本語のレビューにおける共通点や相違点について調査を行うことで、SNS を利用するユーザの国民性などを抽出することを目指す。

2. SNS と感情の関連性

SNS とは

SNS はソーシャル・ネットワーク・サービスを指し、インターネット上のコミュニティサイトのことである。誰もが簡単に情報を発信でき、趣味、職業、居住地域などを同じくする個人同士のコミュニティを容易に構築する事ができる。

総務省が 2021 年 1 月に調査したデータによると、日本の Twitter の使用率は 42.3% だった。日本の人口は 1 億 2,557 万人 (2021 年 1 月現在) なので、日本の Twitter ユーザー数は約 5,310 万人^[3] である。本研究では Twitter は日本を代表する SNS として、ユーザのデータを大量に収集し、分析を行う。

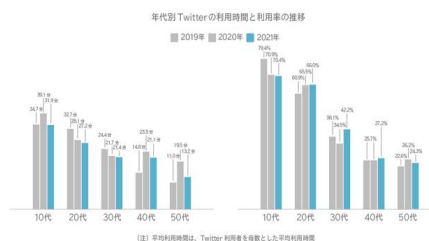


図 1 年代別 Twitter の利用時間と利用率

Twitter 分析の活用事例

SNS 上のデータを分析することができれば、人類社会にも有益な貢献をする可能性がある。大手動画配信サイト Netflix では、Twitter から自社サービスへのコメントを収集する中で、ユーザーの大半が途中で眠ってしまったという問題を発見した。そこから、ユーザーが居眠りしていることを検知し動画を一時停止する機能の開発に至った。健康管理製品を開発している Fitbit では SNS 上から世間の声を抽出して「移動リマインダー」という新しい機能を取り付けた。長時間座り続けているユーザーに警告を鳴らし、運動を促進する機能である^[4]。

Twitter の関連研究

人々 SNS を使って、感情・観点・日常生活などを公開している。2019 年以降、新型コロナウイルス感染症 (COVID-19) の感染拡大により、緊急事態宣言が長期化し、人々の行動が長い間制限されている。医療や病気に関する様々な話題が Twitter で急速に増えている。その中でポジティブ感情と

ネガティブ感情の割合を調べることで、感染者数の指標を推測できる^[5]。また、コロナ予防だけでなく、Twitterに含まれる感情を系統的に分析し、得られた結果を日常的な支援に活用することができれば、社会にとっても有用であると考えられる^[6]。本研究ではTwitterのデータを収集し、より正確な世間の声を把握し、人類社会において有益な貢献を図ることを目的とする。

3. データの収集

実験に必要なデータ情報は、Twitter APIを用いて収集を行う。前章では、Twitter APIの認証により、Twitterサーバからデータを抽出する機能を使用できる。また、プログラミング文を指定することで、Twitter上の異なる関連コンテンツを選択的に抽出することができる。

Twitter API アクセス頻度は15分間隔で、15分以内に180回の命令の実行が可能である(コマンド実行毎に複数のTwitterメッセージを取得可能)^[7]。

| A | B | C | D | E | F | G | H | I | J |
|-----|---------------------|-----------|-------------|---|---|------------------|-----------|-----------|------|
| 検索日 | ID | User name | Screen name | | | favorite retweet | followers | followers | user |
| 1 | 1000000000000000000 | user1 | user1 | | | 0 | 0 | 0 | 0 |
| 2 | 1000000000000000000 | user2 | user2 | | | 0 | 0 | 0 | 0 |
| 3 | 1000000000000000000 | user3 | user3 | | | 0 | 0 | 0 | 0 |
| 4 | 1000000000000000000 | user4 | user4 | | | 0 | 0 | 0 | 0 |
| 5 | 1000000000000000000 | user5 | user5 | | | 0 | 0 | 0 | 0 |
| 6 | 1000000000000000000 | user6 | user6 | | | 0 | 0 | 0 | 0 |
| 7 | 1000000000000000000 | user7 | user7 | | | 0 | 0 | 0 | 0 |
| 8 | 1000000000000000000 | user8 | user8 | | | 0 | 0 | 0 | 0 |
| 9 | 1000000000000000000 | user9 | user9 | | | 0 | 0 | 0 | 0 |
| 10 | 1000000000000000000 | user10 | user10 | | | 0 | 0 | 0 | 0 |
| 11 | 1000000000000000000 | user11 | user11 | | | 0 | 0 | 0 | 0 |
| 12 | 1000000000000000000 | user12 | user12 | | | 0 | 0 | 0 | 0 |
| 13 | 1000000000000000000 | user13 | user13 | | | 0 | 0 | 0 | 0 |
| 14 | 1000000000000000000 | user14 | user14 | | | 0 | 0 | 0 | 0 |
| 15 | 1000000000000000000 | user15 | user15 | | | 0 | 0 | 0 | 0 |
| 16 | 1000000000000000000 | user16 | user16 | | | 0 | 0 | 0 | 0 |
| 17 | 1000000000000000000 | user17 | user17 | | | 0 | 0 | 0 | 0 |
| 18 | 1000000000000000000 | user18 | user18 | | | 0 | 0 | 0 | 0 |
| 19 | 1000000000000000000 | user19 | user19 | | | 0 | 0 | 0 | 0 |
| 20 | 1000000000000000000 | user20 | user20 | | | 0 | 0 | 0 | 0 |
| 21 | 1000000000000000000 | user21 | user21 | | | 0 | 0 | 0 | 0 |
| 22 | 1000000000000000000 | user22 | user22 | | | 0 | 0 | 0 | 0 |
| 23 | 1000000000000000000 | user23 | user23 | | | 0 | 0 | 0 | 0 |
| 24 | 1000000000000000000 | user24 | user24 | | | 0 | 0 | 0 | 0 |
| 25 | 1000000000000000000 | user25 | user25 | | | 0 | 0 | 0 | 0 |
| 26 | 1000000000000000000 | user26 | user26 | | | 0 | 0 | 0 | 0 |
| 27 | 1000000000000000000 | user27 | user27 | | | 0 | 0 | 0 | 0 |
| 28 | 1000000000000000000 | user28 | user28 | | | 0 | 0 | 0 | 0 |
| 29 | 1000000000000000000 | user29 | user29 | | | 0 | 0 | 0 | 0 |
| 30 | 1000000000000000000 | user30 | user30 | | | 0 | 0 | 0 | 0 |
| 31 | 1000000000000000000 | user31 | user31 | | | 0 | 0 | 0 | 0 |
| 32 | 1000000000000000000 | user32 | user32 | | | 0 | 0 | 0 | 0 |
| 33 | 1000000000000000000 | user33 | user33 | | | 0 | 0 | 0 | 0 |
| 34 | 1000000000000000000 | user34 | user34 | | | 0 | 0 | 0 | 0 |
| 35 | 1000000000000000000 | user35 | user35 | | | 0 | 0 | 0 | 0 |
| 36 | 1000000000000000000 | user36 | user36 | | | 0 | 0 | 0 | 0 |
| 37 | 1000000000000000000 | user37 | user37 | | | 0 | 0 | 0 | 0 |
| 38 | 1000000000000000000 | user38 | user38 | | | 0 | 0 | 0 | 0 |
| 39 | 1000000000000000000 | user39 | user39 | | | 0 | 0 | 0 | 0 |
| 40 | 1000000000000000000 | user40 | user40 | | | 0 | 0 | 0 | 0 |
| 41 | 1000000000000000000 | user41 | user41 | | | 0 | 0 | 0 | 0 |
| 42 | 1000000000000000000 | user42 | user42 | | | 0 | 0 | 0 | 0 |
| 43 | 1000000000000000000 | user43 | user43 | | | 0 | 0 | 0 | 0 |
| 44 | 1000000000000000000 | user44 | user44 | | | 0 | 0 | 0 | 0 |
| 45 | 1000000000000000000 | user45 | user45 | | | 0 | 0 | 0 | 0 |
| 46 | 1000000000000000000 | user46 | user46 | | | 0 | 0 | 0 | 0 |
| 47 | 1000000000000000000 | user47 | user47 | | | 0 | 0 | 0 | 0 |
| 48 | 1000000000000000000 | user48 | user48 | | | 0 | 0 | 0 | 0 |
| 49 | 1000000000000000000 | user49 | user49 | | | 0 | 0 | 0 | 0 |
| 50 | 1000000000000000000 | user50 | user50 | | | 0 | 0 | 0 | 0 |

図 2 API より収集したのデータ

4. 機械学習の実行と結果

前章のTwitter APIで収集したテキストデータを感情的に分類し、機械学習のトレーニングを行う。機械学習の結果は混同行列を用いて評価し、学習の評価指標は主に4種類、正解率(Accuracy)、精度(Precision)、検出率(Recall)、F値(F-measure)である。

混同行列

混同行列は予測結果の精度評価を表す標準フォーマットであり、n行n列の行列形式で表される。図3でのTN(True Negative)は本来ネガティブに分類すべきものを正しくネガティブに正

しく分類できたものであり^[8]、FN(False Negative)は本来ポジティブに分類すべきものを誤ってネガティブに分類した件数を示す。FP(False Positive)は本来ネガティブに分類すべきものを誤ってポジティブに分類した件数を示し、TP(True Positive)は本来ポジティブに分類すべきものをポジティブと正しく分類できた件数である。

| | | 予測値 | |
|-------------|------------------------|------------------------|------------------------|
| | | 負事例 (陰性/Negative/0) | 正事例 (陽性/Positive/1) |
| 真 の 値 | 負事例 (陰性/Negative/0) | TN True Negative | FP False Positive |
| | 正事例 (陽性/Positive/1) | FN False Negative | TP True Positive |

図 3 混同行列は二値分類の結果

評価指標

正解率(Accuracy)とは、本来ポジティブに分類すべきものをポジティブに分類し、本来ネガティブに分類すべきものをネガティブに分類できた割合である。すべての予測が正しいサンプルをサンプル総数で割ったものである。

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

精度(Precision)とは、ポジティブに分類されたもののうち、実際にポジティブであったものの割合である。

$$Precision = \frac{TP}{TP + FP}$$

検出率(Recall)とは、本来ポジティブに分類されるべきものを、正しくポジティブに分類できたものの割合である。

$$Recall = \frac{TP}{TP + FN}$$

F値(F-measure)とは、精度(Precision)と検出率(Recall)のバランスを示す指標である。Fの分布区間は0から1の間であり、F値が1に近いほど精度(Precision)と検出率(Recall)のバランスが取れており良い評価となる^[9]。

$$F\text{-measure} = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 * Precision * Recall}{Precision + Recall}$$

実験結果

学習する前のモデルは、テキスト感情を分類判定する際に、テキストを積極的または消極的な2つの感情のうちの1つに分類するしかない。日常生活においても、SNSにおいても、すべての文が感情を持っているわけではないため、判断に感情のない分類を加えることが重要である。本実験では収集したデータを3つの分類処理し、ポジティブとネガティブの2つの分類に基づいて、ナチュラルの感情分類を加えた^[10]。

今回の実験では、3組のデータを用意した。1組あたり3000件のデータである。DATA1はツイッターがランダムに収集したデータであり、DATA2はツイッターで指定されたテーマコンテンツのデータである。DATA3は各アプリ商品評価のユーザ評価データである。BERTモデルに基づいて実験をおこない、その結果を下表に示す。

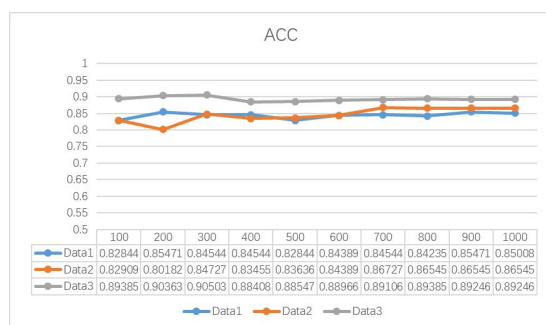


図4 正解率

図4は正解率を示しており、DATA3の正解率は学習過程でいずれもDATA2およびDATA3より優れている。全体的なデータを見ると、商品評価の内容はSNSの内容に対してネット化された用語が少ない。また、商品の評価はSNSの内容よりも2段階に分かれており、「ポジティブ」または「ネガティブ」の判定がしやすいことが分かる。

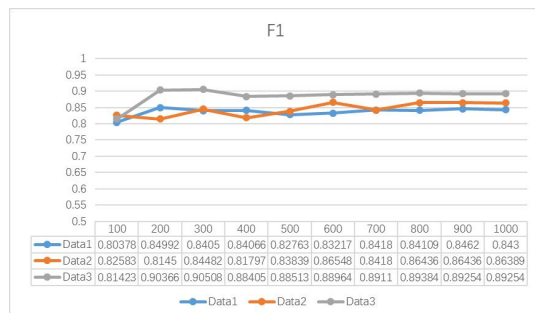


図5 F値

図5に示すデータはF値であり、学習初期の3種類のデータはほぼ同じレベルであった。DATA3に代表される商品評価データは分化しやすいためF値が安定している。テーマが指定されているDATA2では、F値は学習中変動している。DATA2は、ある程度の学習を経て、判定の指標はテーマが指定されていないDATA3を上回っているが、F値は安定していない。テーマを指定したため、オーバーフィットの可能性がある。大量の学習を行うことで、F値の変動は安定してDATA3を上回っている。

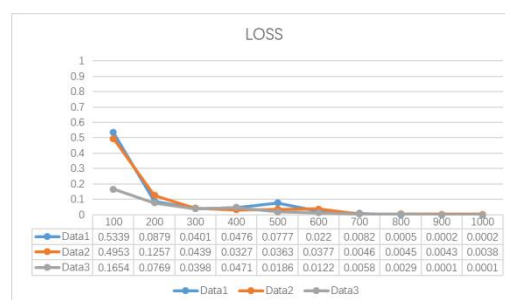


図6 LOSS値

図6のLOSS値は、本実験におけるモデル学習能力の限界を表している。3種類のデータは実験開始後すぐに収束を開始した。さらにDATA1とDATA3は実験完了時に収束を完了し、モデル学習が完了したことを示し、継続しても変動は生じないが、DATA2は完全収束を確定することができず、900~1000の段階では微細な変動が残っている。Epoch^[11]を調整することで、新しい結果が得られる可能性がある。

SNSの内容は、さまざまなネット用語や若者に人気の新興語、世界中の専門用語が含まれているため、機械学習の分野では難関とされてきた。本実験では、SNSのデータおよび商品評価を用いて比較したが、SNS解析にとって完璧なモデルは得られなかったが、その結果は従来の分野の商品評価と比較してそれほど差はなかった。

5. おわりに

本研究では、Twitter上からレビューコメント等を収集し、大量なコメントの中の感情について、効率的に分析する手法の提案を行った。Twitter APIのルールでは最近1週間のデータしか収集す

ることができないため、2019 年や 2020 年のデータと関連付けて分析するには、別の方法でデータを収集する必要がある。また、機械学習に用いたデータは数千件であり、機械学習で得られたモデルの指標をさらに高めるためには、より多くのデータが必要である。今回の実験では感情の判定は著者一名で行っていたが、今後は複数の人が感情判定を行うことにより、学習完了後のモデル予測結果が今回の実験より向上する可能性がある。

参 考 文 献

- [1] Meena Rambocas, João Gama 「Marketing research: The role of sentiment analysis」, FEP – Working Papers, April 2013
- [2] 岡本 哲和, 「日本における候補者のインターネット利用: 2000–2019」, 關西大學法學論集, Vol. 71, No. 3, pp. 547–593, 2021
- [3] 「Twitter (Twitter) 日本のユーザー数・企業の事例・新機能 fleet も紹介」
<https://unique1.co.jp/column/sns_operation/1132/> 【2021/9 更新】
- [4] 「ソーシャルリスニングとは？消費者の生の声を拾う方法と分析ツールを解説」
<<https://keywordmap.jp/academy/what-is-social-listening/#i-6>>
- [5] 峰滝 和典, 「Twitter 上の新型コロナウイルス関連語句の分析—2 度目の緊急事態宣言前後の動向に焦点をあてて—」, 商経学叢, Vol. 67, No. 3, pp. 101–120, 2021
- [6] 工藤 裕登, 砂山 渡, 「ツイート中の感情表現単語の使用割合に基づく性格推定支援」, 2021 年度人工知能学会全国大会 (第 35 回), 2021
- [7] Rate limits Standard API rate limits per window, <<https://developer.twitter.com/ja/docs/basics/rate-limits>>
- [8] 「機械学習の分類における評価指標に関して学ぼう！」, <<https://aiacademy.jp/media/?p=258>>
- [9] 「scikit-learn でクラス分類結果を評価する」, <<https://pythondatascience.plavox.info/scikit-learn/%E5%88%86%E9%A1%9E%E7%B5%90%E6%9E%9C%E3%81%AE%E3%83%A2%E3%83%87%E3%83%AB%E8%A9%95%E4%BE%A1>>
- [10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, 「BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding」, 2018
- [11] 「エポック (epoch) 数とは【機械学習 / Deep Learning】」, <<https://www.st-hakky-blog.com/entry/2017/01/17/165137>>