

甲南大学大学院
自然科学研究科
知能情報学専攻
修士論文 No. 192

ソーシャルメディアにおける多次元感情分類手法

Classifying Multidimensional Sentiment in Social Media using
Machine Learning Methods

2019年3月
秋山 和寛

甲南大学大学院 自然科学研究科

要旨

近年、Twitter に代表されるマイクロブログや Amazon のようなレビューサイトなどのソーシャルメディアの普及に伴い、人々はインターネット上で様々な話題について意見を交わしている。例えば、Twitter では投稿の内容などから共通の趣味を持つ者同士などでフォロー、フォロワー関係によるコミュニティが形成されている。また、レビューサイトでは消費者が自身の購入した商品に対しての意見や感想を投稿しており、これからその商品を購入しようか検討しているユーザなどはこの情報を参考とすることができる。このようにソーシャルメディア上の情報はユーザにとって有益なものが多数ある。しかしながら、これらの情報は膨大であり、その中から自身が必要としているものを探し出すのは大変困難である。そこで本研究では、ソーシャルメディア上の情報の分類において感情が重要な役割を果たす一因であると考え、3種類のソーシャルメディアにおける多次元感情分類手法を提案する。第1に Twitter にはフォロー機能があるが、膨大なアカウントの中からフォローしたいと思うユーザを探し出すのは困難であり、見落としが発生してしまう。そこで、機械学習を用いてツイートの感情を自動で分類し、同じ話題に対するツイートの感情の類似からユーザ間の類似を求める。これにより話題と内容に着目した見落とされがちな潜在的フォロワーの抽出、推薦を行う。第2に、ECサイトなどに投稿されたレビューからは商品に対する購入者の意見を得ることができる。購入者の意見が表れている部分には商品に対する良い点や悪い点、客観的な事実といったポジティブ、ネガティブ、ニュートラルなどの感情を含んでいる場合が多いと考えられる。そのため購入者の意見を得るには感情に基づいた意見の分析、分類が効果的であると考えられるが、レビューの数は膨大であり、個々のレビューの中にも複数の感情が混在していることから、分析にはコストがかかってしまう。そこで、本研究ではレビューを節の単位に分割し、機械学習の1手法である条件付き確率場 (CRF) を用いたレビューの節単位毎の感情分類手法を提案する。これにより、文脈を考慮してより正確にレビュー中の感情を捉え分類を行う。第3に、ショッピングサイトにおける評価の高いレビューの感情分析を行う。評価の高いレビューとレビュー中の感情との関係を分析することで、どのような内容のレビューが他者から参考にされているのかを把握し、参考になるレビューの推薦や作成時の補助となることが期待できる。

Summary

Nowadays, along with the spread of social media such as Twitter and review sites, people discuss their opinions on various contents on the Internet. For example, Twitter users follow other users that have the same interests based on their post. In the case of the review sites, consumers post their opinion about the products that they bought. It is helpful for other consumers to understand the product's performance. These cases indicate that the data on social media is beneficial. However, much data exist on social media. Therefore, it is difficult to extract the data that they need. We propose three methods of classifying multidimensional sentiment in social media. First, we propose a method that extracts latent

followers using a machine learning method. Our proposed latent follower is that people tweet similar content and similar sentiment by using machine learning. In this case, we use Repeated Bisection method to classify a similar sentiment tweet. Second, we propose a clause-level sentiment classification method using Conditional Random Fields (CRF). Usually, we are able to get the consumer's opinion about the product that they bought from reviews. In that case, sentiment classification is useful because the consumer's opinion usually contains some sentiments. However, it takes a high cost because there are a lot of reviews on the Internet and one review contains several sentiments. Then, we propose the clause-level sentiment classification method. Therefore, it becomes easy to analyze consumer's opinion. Finally, we analyze the reviews which receive a higher rate from another user based on multi-dimensional sentiments.

目次

| | | |
|----------|-------------------------------------|-----------|
| 1 | はじめに | 1 |
| 2 | 関連研究 | 3 |
| 2.1 | フォロイー推薦に関する研究 | 3 |
| 2.2 | レビュー分析に関する研究 | 3 |
| 3 | 機械学習を用いた感情抽出に基づく潜在的フォロイー抽出手法 | 5 |
| 3.1 | 話題の類似性 | 5 |
| 3.1.1 | 両ユーザのツイート数の比率 | 5 |
| 3.1.2 | ツイートの凝集性 | 6 |
| 3.1.3 | ガベージツイート | 6 |
| 3.2 | 感情の類似度の算出 | 6 |
| 3.2.1 | 機械学習 | 7 |
| 3.2.2 | 類似度の算出 | 7 |
| 3.3 | 実験 | 8 |
| 3.3.1 | 実験1：多次元感情分類手法に基づくフォロイー推薦の有用性の評価 | 8 |
| 3.3.2 | 実験2：潜在的フォロイー抽出の評価 | 10 |
| 4 | レビューにおけるCRFを用いた節単位毎の感情分類手法 | 13 |
| 4.1 | 投稿文の節単位毎の感情分類 | 13 |
| 4.1.1 | 投稿文の節単位への分割 | 13 |
| 4.1.2 | CRFを用いた感情分類手法 | 14 |
| 4.2 | 実験 | 14 |
| 4.2.1 | 使用データ | 15 |
| 4.2.2 | 実験条件 | 15 |
| 4.2.3 | 実験結果及び考察 | 16 |
| 5 | レビューの感情に基づいた評価の高いレビューの分析 | 20 |
| 5.1 | レビューの感情に基づいた評価の高いレビューの分析 | 20 |
| 5.2 | データセット | 20 |
| 5.3 | 仮説 | 21 |
| 5.4 | 分析結果と考察 | 22 |
| 6 | まとめと今後の課題 | 25 |
| 6.1 | 機械学習を用いた感情抽出に基づく潜在的フォロイー抽出手法 | 25 |
| 6.2 | レビューにおけるCRFを用いた節単位毎の感情分類手法 | 25 |
| 6.3 | レビューの感情に基づいた評価の高いレビューの分析 | 25 |

目 次

| | | |
|---|------------------------------|----|
| 1 | Amazon - 日用品 実験結果 | 19 |
| 2 | Amazon - 家電機器 実験結果 | 19 |

表目次

| | | |
|----|---------------------------------------|----|
| 1 | ツイートの感情値の算出例 | 7 |
| 2 | 教師データの例 | 8 |
| 3 | 被験者の特徴 | 9 |
| 4 | ユーザタイプの分類 | 10 |
| 5 | 多次元感情分類手法に基づくフォロワー推薦の有用性の評価 | 11 |
| 6 | 潜在的フォロワー抽出の評価 | 11 |
| 7 | 極性反転語一覧 | 14 |
| 8 | 節毎の感情ラベルの評価例 | 16 |
| 9 | 感情付与結果 | 17 |
| 10 | 分析に用いるデータセット | 21 |
| 11 | 他者から参考になると評価された回数の一覧 | 22 |
| 12 | ポジティブ, ネガティブ両方の感情が記載されている割合 | 23 |
| 13 | ポジティブの感情が記載されている割合 | 23 |
| 14 | ネガティブの感情が記載されている割合 | 23 |
| 15 | 連続した3つ以上のニュートラルの節が存在している割合 | 24 |

1 はじめに

近年、Twitter¹や Facebook²のようなマイクロブログや Amazon³や楽天⁴のような EC サイトにおけるレビューなどのソーシャルメディアの普及に伴い、人々はインターネット上で様々な話題について意見を交わしている。例えば、Twitter では投稿の内容などから共通の趣味を持つ者同士などでフォロー、フォロワー関係によるコミュニティが形成されている。また、EC サイトのレビューでは消費者が自身の購入した商品に対しての意見や感想を投稿しており、これからその商品を購入しようか検討しているユーザなどはこの情報を参考とすることができる。このようにソーシャルメディア上の情報は有益なものが多数ある。しかしながら、これらの情報は膨大であり、その中から自身が必要としているものを探し出すのは大変困難である。そこで本研究では、ソーシャルメディア上の情報の分類において感情が重要な役割を果たす一因であると考え、3種類のソーシャルメディアにおける多次元感情分類手法を提案する。

まず第1に、機械学習を用いた感情抽出に基づく潜在的フォロワー抽出手法を提案する。代表的なマイクロブログの1つである Twitter では、任意のユーザをフォローすることで容易にそのユーザ（すなわちフォロワー）のツイートから様々な情報を取得することができる。そのため、フォロワーの数を増やせばより多くの情報を取得できるようになるが、Twitter 上には膨大な数のユーザが存在しているため、ユーザの興味と感性に合ったフォロワーを見つけ出すのは容易ではなく、実質的かつ潜在的な不利益を生じている。このような問題を解決するために、Twitter を対象とするフォロワー推薦に関する研究が数多く行われており、主な手法としてネットワーク関係に基づく手法とコンテンツの類似性に基づく手法がある。例えば、前者の例としては、Scott ら [1] がフォロワー・フォロワー関係から現実世界における友人を推定し、推薦する手法を提案している。また、後者の例としては、大村ら [2] がツイートやプロフィールの内容から被推薦ユーザが興味を示している語を抽出し、その語についてつぶやいていたユーザを推薦候補としている。コンテンツの類似性を考慮した推薦手法では、一般にツイートから話題を抽出し、被推薦ユーザと類似した話題についてツイートしているユーザを推薦するが [3], [4], これらの手法では内容の類似性のみ着目しており、その内容に対するユーザの感情までは考慮していない。そのため、ある話題について肯定的なツイートをしているユーザが否定的なツイートをしているユーザに推薦されるといった問題が生じてしまう。したがって、コンテンツの類似性に基づくフォロワー推薦においてはツイートの内容だけではなく、その内容に対するユーザの感情の類似性も考慮することが重要であると考えられる。また、ユーザの感情を単にポジティブ／ネガティブの2つの極性に分類してしまうと、「悲しい」や「嫌い」といった語はどちらもネガティブに分類されることになる。そのため、例えば、ある有名人の引退発表に対して、ファンのユーザが「〇〇が引退するなんて悲しい」とツイートし、その有名人を嫌いなユーザが「〇〇なんて大っ嫌い」とツイートした場合、この2人のユーザはお互いにフォロワー候補として推薦され合うことになる。このようにポジティブ／ネガティブという2つの極性に基づく感情分類では、ユーザ間の感情の類似性を正確に判断するのは難しい場合がある。そこで本研究では、ツイート内容だけでなくその内容に対する感情（「喜・好」、「哀」、「怒・厭」、「安」、「恥」、「怖」、「昂」、「驚」）まで類似しているユーザを潜在的フォロワーと呼び、機械学習による感情抽出を行い、潜在的フォロワーを推薦する手法を提案する。さらに、実験により潜在的

¹Twitter <https://twitter.com/>

²Facebook <https://www.facebook.com/>

³Amazon <https://www.amazon.co.jp/>

⁴楽天 <https://www.rakuten.co.jp/>

フォロワーが推薦できたかどうかを検証する。

第2にレビューにおけるCRFを用いた節単位毎の感情分類手法を提案する。レビューをはじめとして、インターネット上には様々な商品やサービスに対する消費者による投稿文が存在している。中でも、ネガティブな投稿は企業にとって商品やサービスの改善を図る上で有益であると考えられる。しかしながら、投稿文の数は膨大であり、1つの文の中にも複数の感情が混在していることがあるため、人手での分析は負担が大きい。そのため、機械学習によるネガティブ感情の自動分類が効果的であると考えられる。しかしながら投稿文は商品やサービス（以下、ドメイン）によってその特徴が異なるため、機械学習による感情分類では、ドメイン毎に教師データを作成しなければならないとコストがかかる。そこで本研究では、あるドメインの投稿文に加えて、日常生活における様々な不満に関する投稿文（以下、不満投稿）を教師データに加えて、複数ドメインのネガティブ感情分類を行う。これにより単一ドメインのみによる学習との比較、分析を行うことで、マルチドメインによる学習の有用性を測り、消費者投稿文のネガティブ感情分類における不満投稿を用いた転移学習の可能性及び教師データ作成コストの削減の確認を行う。

第3にショッピングサイトにおける評価の高いレビューの感情分析を行う。ネットショッピングにおいてユーザはレビューを参考にして商品を購入する 경우가多数ある。ユーザは商品について書かれたレビューを読むことでその商品についてより深く理解し、購入するかどうかを判断する根拠のひとつとしている。また各レビューには他のユーザがそのレビューが参考になったかどうかの点数が表示されており、この点数から参考に成り得るレビューの目星をつけることがある。しかしながら、必ずしも全てのレビューに点数がついて評価されているわけではなく、参考になる内容であるにも関わらず埋もれてしまっているレビューも多数存在している。一方でユーザの購買意欲には、レビューの中でも特に商品に対する評価や商品に関する事実が影響するということが先行研究から判明している [5]。このことから、他のユーザから参考にされているレビューにはこれらの内容が記載されている事が多いと考えられる。また、商品に対する評価には良い点や悪い点といったポジティブやネガティブの感情が、商品に関する事実にはニュートラルの感情が多く現れると考えられる。以上より、参考にされているレビューとレビュー中の感情との関係を分析することで、どのような内容のレビューが他者から参考にされているのかを把握し、参考になるレビューの推薦や作成時の補助となることが期待できる。本研究では、多くのユーザから参考にされているレビューを「評価の高いレビュー」と呼び、レビューの感情に基づいた評価の高いレビューの分析を行う。

以下、2章で関連研究について述べる。3章では機械学習を用いた感情抽出に基づく潜在的フォロワー抽出手法について、4章でレビューにおけるCRFを用いた節単位毎の感情分類手法について述べる。5章ではレビューの感情に基づいた評価の高いレビューの分析。6章ではまとめと今後の課題について述べる。

2 関連研究

2.1 フォロワー推薦に関する研究

Twitter に代表される SNS の普及に伴い、フォロワー推薦に関する研究が数多く行われている。フォロワー推薦の手法はユーザ同士のつながりを考慮したネットワークに基づく推薦手法と、投稿内容の類似性に着目したコンテンツの類似性に基づく推薦手法の 2 つに大きく分けられる。また、一方で SVM による感情分析に関する研究も盛んに行われている。以下、これら 2 種類の手法及び感情分析について説明する。

ネットワークに基づく推薦手法

Armentano ら [6] は、フォロワー・フォロワー関係に代表される Twitter ネットワークトポロジーを探索するアルゴリズムに基づいたフォロワー推薦システムを提案している。北村ら [7] は、リプライ等のユーザ間関与に基づく 2-hop ユーザ推薦モデルを提案している。これらの研究では、ユーザ間のフォロワー・フォロワー関係を考慮してフォロワー候補を推薦しているため、趣味や興味のあることについてツイートしていないユーザでもネットワーク関係から興味のあるユーザを推薦することが可能となっている。しかしこの手法はユーザのネットワーク関係に依存しているため、閲覧ユーザとフォロワー・フォロワーのフォロー傾向が異なる場合、閲覧ユーザにとって興味のないユーザが推薦されてしまうことがある。これに対し、本研究では閲覧ユーザの投稿した話題と類似しているユーザを推薦しているため、興味のあるユーザを推薦することができる。

コンテンツの類似性に基づく推薦手法

黒柳ら [8] は共通の著名人をフォローしているということは、その著名人のファン、若しくは何かしらの興味、関心がある可能性が高いと考え、著名人を共通にフォローしているユーザをフォロー候補として提示している。桑原ら [9] は投稿者のメッセージから共通の話題に対して同じような感情を持つ投稿者の推薦を行っている。これらの手法では閲覧ユーザと興味・関心が似ているユーザをフォロワー候補として推薦しているが、その興味や関心に対する両ユーザの感情まで同じであるとは限らないため、推薦候補としては不適切なユーザを推薦する可能性がある。これに対し本研究では話題に対する感情の類似性まで考慮している点が異なる。

2.2 レビュー分析に関する研究

“Amazon” や “TripAdvisor”⁵ といったレビューサイトの普及に伴い、投稿者の意見を分析するためにレビューのトピック抽出に関する研究が数多く行われている。また、同様にレビューの感情分析に関する研究も行われている。以下でそれらについて説明する。

レビューのトピック抽出

Rustamov ら [10] は主観的な部分には意見が書かれているという仮定のもと、2 つの機械学習によるアプローチから文レベルでの主観部分抽出を行っている。Paltoglou ら [11] は主観性及び極性の 2 つの値で各文を特徴付け、機械学習による分類を行っている。田村ら [12] はユーザ及び商品のトピックに加え、その商品に付けられた評点を加えたトピックモデルの構築を提案している。これらの研究では投稿文全体や投稿文中の各文を対象としているのに対し、本研究では節を対象とした感情推定を行っている点が異なる。

⁵<https://www.tripadvisor.jp/>

レビューの感情分析

レビューを分析する上で、本文中の内容に対する投稿者の感情分析を行うことはとても重要である。鈴木ら [13] はラベル付きデータに加えラベルなしデータを用いた学習を行うことで単体で感情を持つ語のみならず、形容詞のような前後の文脈を考慮しなければ感情がわからない語にも対応した評価表現抽出手法を提案している。Lin ら [14] はレビュー本文の背景にある感情を仮定しトピックモデルを構築している。Choi ら [15] は単語間の相互作用による極性の生成に着目し、否定語が及ぼす影響を考慮した表現レベルでの感情分析を行っている。これらに対し、本研究では要望部分を抽出し、それに対し感情分析を行うという点が異なる。

3 機械学習を用いた感情抽出に基づく潜在的フォロワー抽出手法

本章では、機械学習を用いてツイートの感情抽出を行うことで、ツイートの内容とそれに対する投稿者の感情の類似に着目した潜在的フォロワー抽出手法を提案する。我々の提案する潜在的フォロワー抽出手法は、ユーザ同士のコンテンツの類似性と感情の類似性に基づいている。具体的には、まずコンテンツ（すなわちツイートの話題）が類似するユーザを決定し、次にそのユーザの中から類似したコンテンツ（すなわち共通の話題）に対して感情も類似しているユーザを探し出し潜在的フォロワーとして推薦する。尚、本論文では推薦を受けるユーザのことを「閲覧ユーザ」と呼ぶ。また本論文では、実験の都合上、フォロワーの推薦候補となるユーザを閲覧ユーザのフォロワーのフォロワーかつ閲覧ユーザの既存フォロワーでないユーザとする。このフォロワーのフォロワーを本論文では「ffユーザ」と呼ぶ。提案手法の流れを以下に示す。

- 1 閲覧ユーザの ffユーザを 200 名ランダムに取得する。
- 2 以下の手順を 200 名の ffユーザに繰り返す
 - 2.1 閲覧ユーザとある ffユーザの最新のツイートを 200 件ずつ収集する。
 - 2.2 収集した 400 件のツイートから両ユーザに共通の話題を抽出する。
 - 2.3 機械学習を用いて各共通の話題に含まれるツイートの感情を分類する。
 - 2.4 共通の話題内のツイートの感情の類似度を求め、類似度が閾値以上の場合、該当の ffユーザをフォロワー候補として推薦する。

以下、「共通の話題の抽出」、「機械学習によるツイートの感情分類」及び「感情の類似度の計算」について説明する。

3.1 話題の類似性

本研究では話題の類似性として、閲覧ユーザと ffユーザの共通の話題をクラスタリングにより決定する。具体的には、閲覧ユーザと ffユーザの最新のツイートをそれぞれ 200 件ずつ、合計 400 件収集する。次に、収集したツイート群から共通の話題を抽出するために、Repeated Bisection 法 [16] を用いてクラスタリングする。また、クラスタリングツールには Repeated Bisection 法が実装されたツールである bayon[17] を用いる。その結果得られる各クラスタには話題を示す名詞が複数存在しうるが、本研究ではクラスタの中心ベクトルに最も近い名詞をそのクラスタの話題とする。このとき、クラスタの分割数については我々の予備実験の結果より、10 クラスタとする。これら分割された各クラスタから共通の話題を抽出するに際し、クラスタから両ユーザのツイートの比率、ツイートの凝集性、ツイートのガベージ性の 3 点に着目することにより、類似性の高い話題を共通話題として抽出する。

3.1.1 両ユーザのツイート数の比率

あるクラスタ内に閲覧ユーザと推薦候補ユーザ両名のツイートが存在している場合、そのクラスタは両者にとっての共通の話題のクラスタとなる可能性がある。しかしながら、どちらかのツイートがそのクラスタ内の大半を占めているような場合もあれば両者のツイート数が同程度の場合もある。この場合、後者のほうが共通の話題としては相応しいと考え、以下の式を用

いる。具体的には、 i 番目のクラスタにおける両者のツイート数の比率 R_i を求め、その値が閾値 T_R 以下であるクラスタの話題を両者にとって共通の話題とする。

$$R_i = \frac{|X_i - Y_i|}{|X_i| + |Y_i|}$$

ここで、 $|X_i|$ は i 番目のクラスタに属するユーザ X のツイート数を示し、 $|Y_i|$ は i 番目のクラスタに属するユーザ Y のツイート数を示す。

3.1.2 ツイートの凝集性

クラスタリングに用いた Repeated Bisection 法はハードクラスタリングであるため、クラスタリングの対象となったツイートは必ずいずれかのクラスタに分類される。このため、相互に関連性のない話題を含むツイートで構成されたクラスタが存在している場合がある。本研究ではこのようなクラスタをガベージクラスタと呼ぶ。ガベージクラスタの話題については、たとえばクラスタ内の両者のツイート数が同じだとしても共通の話題とは言い難いため、取り除く必要がある。この際、ガベージクラスタ内のツイート同士の話題の関連性が低いという点を利用する。具体的には、 i 番目のクラスタ C_i のセントロイド c_i とそのクラスタに含まれるツイート x_j のコサイン類似度をツイートごとに求める。そして、その平方和をクラスタ C_i の凝集性 A_i と定義し、その値が閾値 T_A 未満のクラスタをガベージクラスタとして取り除く。

$$A_i = \sum_{x_j \in C_i} \left(\frac{x_j \cdot c_i}{|x_j| |c_i|} \right)^2$$

3.1.3 ガベージツイート

あるクラスタ内のツイートの凝集性が閾値 T_A 以上であったとしても、そのクラスタの話題に対して関連性の低いツイートはクラスタ内に存在し得る。そこで本研究ではそのようなツイートをガベージツイートと呼び、該当クラスタ内から取り除く。具体的には、あるクラスタ中の各ツイートのクラスタへの所属度を取得し、その値が閾値 T_G 以下であるツイートをガベージツイートとして取り除く。尚、所属度とはあるツイートとそのツイートが存在しているクラスタの中心ベクトルとのコサイン類似度の値を指す。

以上より、両者のツイートの比率 R_i が閾値 T_R 以下であり、かつツイートの凝集性 A_i が閾値 T_A 以上であるクラスタからガベージツイートを取り除いたものを「共通話題クラスタ」と呼び、共通話題クラスタの中心ベクトルに最も近い名詞を共通の話題として扱うことにする。尚、それぞれの閾値は我々の予備実験により、 $T_R = 0.25$, $T_A = 0.60$, $T_G = 0.25$ とする。

3.2 感情の類似度の算出

これまで我々はツイートの感情抽出に我々の作成した感情語辞書を使用し、ツイート文中の単語ごとの感情値の和からユーザ同士の感情の類似度を算出してきた。ツイートの感情値の算出例を表 1 に示す。しかしながら、実験より感情値算出の精度が十分であるとは言い難いことがわかった。そこで本論文では感情の抽出に機械学習を用い、抽出された感情の類似度から潜在的フォロイーを決定する。ここで、本手法で使用する 8 軸の感情軸は我々の提案してきた「喜・好」、「哀」、「怒・厭」、「安」、「恥」、「怖」、「昂」、「驚」[18] を用いる。

表 1: ツイートの感情値の算出例

| ツイート「梅雨はジメジメして嫌い」 | | | | | | | | |
|-------------------|-------|---|-------|---|---|------|---|---|
| 感情語 | 喜・好 | 安 | 昂 | 哀 | 怖 | 怒・厭 | 驚 | 恥 |
| 梅雨 | 0.14 | 0 | 0.248 | 0 | 0 | 0.09 | 0 | 0 |
| ジメジメ | 0.276 | 0 | 0 | 0 | 0 | 1.29 | 0 | 0 |
| 嫌い | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 合計 | 0.416 | 0 | 0.248 | 0 | 0 | 2.38 | 0 | 0 |

3.2.1 機械学習

機械学習には SVM のライブラリである LibSVM[19] を使用する。SVM のカーネルは RBF カーネルを用い、コストパラメータ c は 2.0, カーネルパラメータ g は 0.0315 とする。使用する教師データと素性は以下の通りである。

教師データ

使用する教師データにはアンケートにより感情ごとに分類されたツイートを用いる。具体的には、まず無作為に取得した 25 件のツイートに対し被験者 10 名が中村らの 10 の感情軸 [20] に感情を持たない軸である「その他」を加えた 11 軸の各感情がそれぞれの程度当てはまっているかを 5 段階で評価した。これを 1 セットとし、全部で 114 セット行った。結果、2,850 件のツイートの評価値を取得した。そして、評価データ内で評価値が 3 以下のものは 0 に置き換え、4 以上のものは 1 に置き換える。そして、それぞれの感情軸毎に評価値が最も高い人と最も低い人の評価値を削除し、残った 8 名の被験者の評価値の和を求め、ある軸の和が他の軸の和に対し 2 倍以上大きければ、その軸をツイートの感情とする。ただし、我々のこれまでの研究 [18] から「安」と「昂」の感情軸は単独で存在することがなく、他の感情軸を補助する役割があることが判明している。このことから、「安」「昂」の 2 軸に関しては、他の軸よりも評価値の和が大きければ該当ツイートの感情とする。そうして取得したツイートは 2,002 件となり、この 2,002 件のツイートを教師データとする。教師データの一例を表 2 に示す。

素性

素性は教師データに対し形態素解析を行い得られた名詞、動詞、形容詞を対象とし、各語がツイートの含まれていれば 1, 含まれていなければ 0 とする。

3.2.2 類似度の算出

感情の類似度の算出には各共通話題クラスタ内の両ユーザのツイートの感情の割合を基にする。ある共通話題クラスタ i における感情の類似度の計算には、2 つの確率分布間の差異を測る尺度である JS ダイバージェンスを用いる。

$$JS_i = \frac{1}{2}KL(P_i||M_i) + \frac{1}{2}KL(Q_i||M_i)$$

$$KL(P_i||M_i) = \sum_{j=1}^8 P_{ij} \log(2) \frac{P_{ij}}{M_{ij}}$$

表 2: 教師データの例

| 感情軸 | ツイート |
|-----|---------------------------------------|
| 喜 | 一緒に面接した人がここはこうしたほうがいってアドバイスくれて嬉しい |
| 好 | 最近気がついたんだけど、私ストライプがけっこう好きみたい |
| 哀 | 対抗戦参加はしてるけど戦力になれなくてすごく悲しい |
| 怒 | まじ洒落ならんわ腹立つー。 どない騒いどんねんカス |
| 厭 | 自分を大切にできない奴は、大嫌いだよ |
| 安 | 大阪に帰ってきて一週間たったけど、実家って安心するね |
| 恥 | この間買ったサンダルが歩く度にキュッキュウいってる。ちょっと恥ずかしいな。 |
| 怖 | 夜一人でふらふらするのやめよ。ちょっと怖い |
| 昂 | 明日一限体育!! しかもバスケできる!!! 興奮しすぎて寝れないぞ!! |
| 驚 | 衝撃的すぎる事実を聞かされてどー反応していいかわからへん |

$$KL(Q_i || M_i) = \sum_{j=1}^8 Q_{ij} \log(2) \frac{Q_{ij}}{M_{ij}}$$

$$M_{ij} = \frac{1}{2}(P_{ij} + Q_{ij})$$

ここで j は 8 軸の感情軸を示し, P_{ij} は共通話題クラスタ i 内における閲覧ユーザの感情軸 j の感情のツイートの割合を表し, Q_{ij} は f ユーザの同クラスタ i 内における感情軸 j の感情のツイートの割合を表す. また, M_{ij} は共通話題クラスタ i における P_{ij} と Q_{ij} の平均の値である. 上記の式より, 両ユーザのある共通話題クラスタ i の感情の割合の類似度 JS_i を求める.

次に以下の式を用いて, 閲覧ユーザ a と f ユーザ b の各共通話題クラスタの感情の類似度の和を求める.

$$DS_{ab} = \sum_{i=1}^k JS_i$$

ここで k は共通話題クラスタの数とする. こうして, 各共通話題クラスタに対する感情の割合の類似度の総和 DS_{ab} を求め, その値が閾値 T_J 未満であるとき, f ユーザを閲覧ユーザに推薦する. なお, 閾値 T_J は我々の予備実験により, 0.43 とする.

3.3 実験

3.3.1 実験 1: 多次元感情分類手法に基づくフォロワー推薦の有用性の評価

感情語辞書を用いて算出した感情値を基に推薦を決定する従来手法 [21] をベースラインとして, 提案手法である機械学習により分類された感情から類似度を算出する手法と比較することで, 提案手法の有用性を測った. さらに, 被験者をタイプ毎に分類し比較することにより, 提案手法がどのようなタイプのユーザに有用であるかの実験を行った.

実験条件

被験者は 20 代の男性 9 名, 女性 1 名の計 10 名を対象とする. 10 名の被験者のツイート及びフォローの傾向を 3 に示す. 実験手順は, 10 名の被験者に自身の f ユーザ 200 名分のツイートを見てもらい, 各 f ユーザごとに実際にフォローしたいかどうかの 2 値分類を行った. そして, 被

験者がフォローしたいと評価した n ユーザを正解データとし、提案手法、ベースラインそれぞれの手法による推薦結果との適合率、再現率、F 値を求めた。

表 3: 被験者の特徴

| 被験者 | ツイート傾向 | フォロー傾向 |
|--------|----------------------------|--------------------------------|
| ユーザ 1 | 日常生活に関することや フォロワーへのリプライ | 現実の友人, 知人 |
| ユーザ 2 | 日常生活に関すること | 現実の友人, 知人, 有益な情報を発信しているユーザ |
| ユーザ 3 | 日常生活に関すること | 現実の友人, 知人 |
| ユーザ 4 | 日常生活に関することや フォロワーへのリプライ | 現実の友人, 知人 |
| ユーザ 5 | 日常生活に関すること | 現実の友人, 知人, 有益な情報を発信しているユーザ |
| ユーザ 6 | 趣味のゲームに関すること | ゲームに関するユーザ |
| ユーザ 7 | 趣味のゲームに関すること | ゲームに関するユーザや 好みの画像を投稿しているユーザ |
| ユーザ 8 | 政治に関すること | 政治や趣味のゲームに関するユーザ |
| ユーザ 9 | 趣味のゲームに関するユーザ | ゲームに関するユーザ |
| ユーザ 10 | 趣味のゲームに関するユーザ | ゲームに関するユーザ |

10名の被験者はツイートしている内容の傾向とフォローしているユーザの傾向から2種類のタイプに分類をする。ユーザ1-5の5名の被験者に関しては日常的な事に関するツイートをよく投稿し、フォロワーは主に現実での友人・知人で構成されており、Type-Aとする。それに対し、ユーザ6-10の5名の被験者は趣味に関するツイートが多く、ネット上の友人や知人をよくフォローしており、Type-Bとする。ユーザタイプの分類に関して表4に示す。

結果と考察

実験結果を表5に示す。表5より、適合率とF値については全般的に提案手法がベースラインと比べて相対的には高い値となっているが、提案手法の適合率、F値とも値自体はあまり高いものではない。これは感情分類の際にはツイートの持つ感情をどれか1つに限定して分類しており、ツイートが複数の感情を併せ持っている場合や感情を持っていない場合を考慮できておらず、精度が下がってしまったためと考えられる。また、再現率に関してはベースラインが非常に高い値を示している。これは、ベースラインでは、 n ユーザ200名の多くが推薦されるという結果になっているためである。以上のことから、提案手法である機械学習を用いた手法のほうが有用であるといえる。しかしながら、提案手法の適合率、再現率、F値のいずれにおいても十分高いとは言えない。教師データ数を増やす、複数感情や無感情への分類にも対応していくといったことを行い、精度を改善していく必要があることがわかった。

また、ユーザのタイプ毎に実験結果を比較すると、どの値においてもType-BのほうがType-Aよりも高い値を示している。これはType-Aのユーザはツイートの内容ではなく、友人や知人といったユーザ自身の情報に着目しフォローの基準としているユーザであるためだと考えられる。また、Type-Aのユーザには日常で起こった嫌な事や腹が立つ事など、ネガティブな事柄につい

表 4: ユーザタイプの分類

| フォロー傾向 | ツイート傾向 | |
|------------|--------------------------------------|---------------------------------------|
| | 日常に関するツイート | 趣味に関するツイート |
| 現実での友人・知人 | ユーザ1 ユーザ2 ユーザ3 ユーザ4 ユーザ5 | - |
| ネット上の友人・知人 | - | ユーザ6 ユーザ7 ユーザ8 ユーザ9 ユーザ10 |

てよく投稿しているという傾向が見られる。しかしながら、同じようにネガティブなことをツイートしているユーザをフォローしたいというわけではないためであると考えられる。Type-Bのユーザについては、ユーザ6とユーザ8がType-Bのユーザの中では特に低い値となっているが、これは、他のType-Bのユーザが映画やゲーム、アニメといった複数の趣味に関してツイートしているのに対し、ユーザ6はある1つのゲームのことについてのみツイートしているためである。また、ユーザ8は主に政治に関することをツイートしており、「怒」の感情を含んだツイートが多いが、実際にユーザ8がフォローしているフォロワーのツイート内容はネガティブではなくポジティブなものが多い。このように自身のツイートの感情とフォロワーの感情の傾向に差があるため、値が低くなったと考えられる。このことから提案手法はType-Bのユーザに適している事がわかった。さらにType-Bのユーザは、複数の話題のツイートをしており、自身のツイートの感情傾向とフォロワーのツイートの感情傾向が類似しているユーザに特に適していると言える。

3.3.2 実験2：潜在的フォロワー抽出の評価

提案手法では、閲覧ユーザが気がついていないが、提示されることによりフォローしたいと気付くような潜在的フォロワーを抽出することを目的としている。そこで、我々は被験者である閲覧ユーザがffユーザから選択したフォロワー候補である実験1の正解データではなく、我々の提案するシステムにより推薦されたフォロワー候補から閲覧ユーザがフォローしたいと選択した中には存在しているユーザが潜在的フォロワーであると考え、実験を行った。

実験条件

被験者は実験1の結果より、提案手法に最も適していると考えられるユーザ7、ユーザ9、ユーザ10の3名である。使用するデータセットは実験1で用いたデータセットからシステムにより推薦された潜在的フォロワー候補とする。実験の流れは以下のとおりである。

1. 被験者に対し、ffユーザ200名の中から提案システムが推薦した潜在的フォロワーの候補を提示する。

表 5: 多次元感情分類手法に基づくフォロイー推薦の有用性の評価

| ユーザ タイプ | 被験者 | 適合率 | | 再現率 | | F 値 | |
|------------|--------|------------|-------|------------|-------|------------|-------|
| | | ベース ライン | 提案手法 | ベース ライン | 提案手法 | ベース ライン | 提案手法 |
| Type-A | ユーザ 1 | 0.098 | 0.116 | 0.950 | 0.571 | 0.178 | 0.193 |
| | ユーザ 2 | 0.057 | 0.047 | 0.917 | 0.400 | 0.108 | 0.083 |
| | ユーザ 3 | 0.087 | 0.091 | 0.889 | 0.353 | 0.158 | 0.145 |
| | ユーザ 4 | 0.114 | 0.131 | 0.826 | 0.550 | 0.201 | 0.211 |
| | ユーザ 5 | 0.078 | 0.133 | 0.882 | 0.407 | 0.143 | 0.200 |
| Type-B | ユーザ 6 | 0.095 | 0.157 | 0.900 | 0.684 | 0.171 | 0.255 |
| | ユーザ 7 | 0.143 | 0.186 | 1.000 | 0.643 | 0.250 | 0.288 |
| | ユーザ 8 | 0.071 | 0.055 | 0.933 | 0.429 | 0.131 | 0.098 |
| | ユーザ 9 | 0.169 | 0.186 | 0.941 | 0.679 | 0.287 | 0.292 |
| | ユーザ 10 | 0.347 | 0.394 | 0.985 | 0.597 | 0.513 | 0.474 |

表 6: 潜在的フォロイー抽出の評価

| 被験者 | 正解データ数 (実験 1) | 正解データ数 (実験 2) | フォロイー候補の 差分 | 適合率 (実験 1) | 適合率 (実験 2) |
|--------|------------------|------------------|----------------|---------------|---------------|
| ユーザ 7 | 18 | 17 | 5 | 0.186 | 0.175 |
| ユーザ 9 | 18 | 46 | 32 | 0.186 | 0.451 |
| ユーザ 10 | 37 | 48 | 16 | 0.394 | 0.511 |

2. 被験者は提示された潜在的フォロイーの候補をフォローしたいかどうかを判断する。
3. 実験 1 の結果と (2) の結果を比較し、実験 1 から (2) にかけて増えた正解を潜在的フォロイーとする。

結果と考察

実験結果を表 6 に示す。ユーザ 9 は適合率が大きく上がっており、提案システムにより推薦された潜在的フォロイーの候補を提示することで、ユーザにとって潜在的なフォロイーに気がつくことができたと考える。ユーザ 10 についても同様に適合率が上がっているが、ユーザ 9 と比べて上昇率は低い。これはユーザ 10 が他の人をフォローすることに積極的なユーザであり、実験 2 の時点でかなりの数の 5 ユーザをフォローしたいと選択しているため、実験 1 と実験 2 の間でフォローしたいと思う人数がそこまで増えなかったと考えられる。一方、ユーザ 7 の上昇率は低い結果になっている。これはユーザ 7 がフォローするかどうかを決める際に、対象ユーザが投稿している画像を判断基準の一つにしているが、現在の我々の提案手法では画像を考慮していない。このため、提案システムはユーザ 7 に対し、文章はユーザ 7 の好みに合っているが、ユーザ 7 が好みでない画像を投稿している 5 ユーザを潜在的フォロイーとして提示した。このことがユーザ 7 の結果が芳しくなかった原因であると考えられる。

以上の結果をまとめると、実験 1 と実験 2 の結果より提案手法は潜在的なフォロイーを抽出することができていると言える。しかしながら、ツイート本文から話題を抽出するという提案

手法の性質上，画像に対しては内容を考慮することができていない．このことから，今後フォロワー推薦の際に画像も考慮する必要があると考えられる．

4 レビューにおける CRF を用いた節単位毎の感情分類手法

現在、インターネット上には消費者による商品やサービスに対する感想や意見が多数投稿されている。投稿された消費者の意見の中には、“汚れが綺麗に落ちてくれて、とても助かっています。”というような商品やサービスの良い点について述べているポジティブなものがある。また一方で“粘着力が弱くすぐに剥がれてしまいます。”といったネガティブなものも存在している。これら投稿文の中でも、ネガティブな投稿は商品やサービスに対する消費者の不満点や改善してほしい点が書かれていることが多い。そのため、このようなネガティブな情報は企業にとって商品やサービスの改善、問題の解決を図る上で特に有益であるといえる。

しかしながら、インターネット上に存在する投稿文の数は膨大である。さらに、“美味しいんだけど、ちょっと量が物足りない。”という投稿文のように、1つの文の中にも異なる感情が混在していることがある。このため、人手でのネガティブ感情分類は負担が大きい。これまでの研究において、レビューサイトの投稿文に含まれる3つの感情(ポジティブ、ネガティブ、ニュートラル)の中で、ネガティブな投稿文は最も占める割合が少ない。そのため、他の感情と比べて教師データとしての情報量に差が出てしまうという問題があることがわかった。また、投稿文は商品やサービス(以下、ドメイン)によってその特徴が異なる。そのため、機械学習による感情分類ではドメイン毎に教師データを作成する必要があるが、コストが掛かってしまう。そこで本論文では、あるドメインの投稿文に加えて、日常生活における様々な不満に関する投稿文(以下、不満投稿)を教師データに用いる。そして、ネガティブな表現を補完した教師データが異なるドメインの投稿文におけるネガティブ感情分類に対応できるかを確認する。この際、単一ドメインのみによる学習や不満投稿のみによる学習との比較を行うことで、異なるドメインの投稿文に対するネガティブ感情分類において、不満投稿を用いたことによる転移学習の可能性及び教師データ作成のコスト削減の確認を行う。

4.1 投稿文の節単位毎の感情分類

4.1.1 投稿文の節単位への分割

投稿文の感情分類を行う際に、1つの文の中にも複数の感情が存在していることがある。そのため投稿文の感情を正確に分類するためには、“文”より小さな単位を対象とすることが必要であると考えられる。本研究では、文の構造において語や句の集合により1つの意味を形作っている“節”を感情分類の対象とする。

投稿文の各節に対しての感情分類を行う前処理として、各投稿文を節の単位に分割する。具体的には、ある投稿文に対して、まず形態素解析システムにより投稿文を形態素に分解する。次にこの出力結果に対して、構文解析を行う。解析結果から投稿文内の節情報を取得し、それに基づいて投稿文を節単位に分割する。この際分割された節の文字数が5文字以下であった場合には、その節を次の節と統合する。また、その節が最後の節であった場合には1つ前の節と統合する。本論文では形態素解析にJuman⁶を、構文解析に黒橋らの開発した日本語構文・格・照応解析システム(KNP)⁷を用いる。

例えば、「ゴム手袋の中に水が入ると気持ち悪いし、なかなか乾かない。けどひっくり返すのが大変」という投稿文の場合、解析結果から「ゴム手袋の中に水が入ると」「気持ち悪いし」

⁶<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

⁷<http://nlp.ist.i.kyoto-u.ac.jp/?KNP>

「なかなか乾かない」「だけどひっくり返すのが大変」という節に区切り，これらの節を用いる。

4.1.2 CRF を用いた感情分類手法

本論文中で感情分類手法として使用する CRF は構造学習により系列ラベリングを解く手法であり，学習の際に周辺の節の情報も併せて学習することで，文法や文脈などの節同士の関係性も考慮した感情分類を行う．学習には CRF の実装である CRFsuite[22] を用い，学習のアルゴリズムには凸最適化問題を効率的に解くのに適している L-BFGS 法 (L-BFGS) を用いる．パラメータには L1 正則化の係数として 1.0，L2 正則化の係数として 0.001 を設定する．また学習の際には，ある節に加えてその前後 2 節までの情報も併せて学習する．

素性

学習に用いる素性には，ある節の中に含まれる各語の原形と品詞，そして極性値及び極性反転語の数をを用いる．ここで本研究における極性反転語とは「けれども」や「それなのに」など，後ろの節の感情に影響を与える可能性のある語のことを指す．本研究では国立国語研究所⁸より公開されている分類語彙表増補改訂版データベース [23] に登録されている「反対」の意味を持つ接続詞，接続助詞 39 語を極性反転語とする．極性反転語の一覧を表 7 に示す．また，語の極性値に関しては，高村ら [24] が作成した単語感情極性対応表の値を基に求める．具体的には，ある語の極性値 x を単語感情極性対応表の感情極性値 n に応じて以下のように決定する．

$$x = \begin{cases} 2 & (n \geq 0.5) \\ 1 & (0.5 > n \geq 0) \\ -1 & (0 > n \geq -0.5) \\ -2 & (n < -0.5) \end{cases}$$

表 7: 極性反転語一覧

| | | | | |
|------|--------|---------|----------|-------|
| しかし | しかしながら | さりながら | けれども | けれど |
| だけども | だけれど | だけども | だけど | けど |
| が | だが | しかれども | されど | ところが |
| しかるに | でも | それでも | さりとて | そのくせ |
| くせに | それでいて | にもかかわらず | なのに | それなのに |
| いえども | とはいえ | とほいうものの | そうはいうものの | その代わり |
| かえって | 翻って | 反面 | 一面 | 一方 |
| 他方 | 反対に | 逆に | それに対して | |

4.2 実験

機械学習を用いた投稿文の感情分類においては分類データと同じドメインの教師データが必要となる．しかしながら，分類を行うドメイン毎に教師データを用意しているのはコストが嵩ん

⁸<https://www.ninjal.ac.jp/>

でしまう。また、本研究においては商品レビューの文中からネガティブな感情を分類することを目的としているが、商品のレビュー文内においてネガティブな感情が占める割合は小さい。そのため教師データ中におけるネガティブなデータの情報も他の感情と比べて少なくなってしまう。そこで本実験では教師データに商品のレビュー文のみならず、ネガティブな感情が多く含まれている、日常生活における不満に関する投稿文を用いてネガティブな感情の分類精度を測る。また、実際に企業が商品レビューからネガティブな情報を抽出する際には、できるだけ取りこぼしを少なくすることが肝要であると考え、本研究では特に再現率に注目して実験を行う。

尚、本実験では異なるドメインの投稿文として、Amazonの「日用品」、「家電機器」カテゴリのレビューをそれぞれ用い、各カテゴリをそれぞれ1つのドメインとする。また日常生活における不満に関する投稿文として、不満買取センター⁹のサービスへ寄せられた幅広いトピックに関するユーザの不満投稿を用いる。

4.2.1 使用データ

実験にはクラウドソーシングにより節毎に感情を付与された投稿文を用いる。具体的には、Amazonの「日用品」、「家電機器」カテゴリの商品レビューを各1,500件、不満に関する投稿文として、不満買取センターへ投稿された投稿文からジャンルを問わずランダムに取得した1,500件を用いる。また、Amazonの「日用品」カテゴリのレビューとしては「トイレトーパー」、「洗剤」、「防虫剤」の商品レビューを各500件、「家電機器」カテゴリのレビューとしては「ノートPC」、「ロボット掃除機」、「スマートフォン」の商品レビューを各500件用いる。クラウドソーシングの具体的な内容としては、上記の各投稿文をKNPにより分割した計34,070件の節それぞれに対して10人の作業者が感情の評価を行う。作業者は各節を読み、ポジティブ、ネガティブ、ニュートラルの3種の感情の中から適していると判断したものを選択する。尚、作業者には判断の指標として、各感情の定義及び例文をあらかじめ提示した。そして、10人の作業者の内、過半数の評価が特定の感情に集まった場合には、その感情を該当の節の感情とする。いずれの感情にも過半数の評価が集中しなかった場合には、その節には明確な感情が表れていないと判断しニュートラルの感情とする。感情付与タスクの例を表8に、結果及び各ドメインに対する作業者のカッパ係数を表9に示す。

表9より、不満買取センターのデータに関しては日常生活における不満が投稿されているため、ネガティブの占める割合が多く、反対にポジティブなデータはとても少なくなっている。Amazonの投稿文内における感情の割合としては、いずれの商品でもニュートラルが最も多く、次いでポジティブ、そしてネガティブは占める割合が最も少ない。また日用品、家電機器のデータ数に大きな差が見られるが、これは家電機器カテゴリの商品のレビューは日用品のレビューと比べて1文当たりの文量が多い傾向があり、それに伴い分割された節の数も多くなったためである。

4.2.2 実験条件

教師データの違いによる分類精度の変化を確認するため、異なるドメインであるAmazonの「日用品」、「家電機器」の投稿文それぞれに対して以下の実験を行った。

実験1 同一ドメインの投稿文のみによる学習

⁹<http://fumankaitori.com/>

表 8: 節毎の感情ラベルの評価例

| 節 | ラベル |
|---|--------------------------|
| 「この掃除機は性能が良いのだが、音がうるさいので、音が静かなモーターにしてほしい」 | |
| この掃除機は性能が良いのだが、 音がうるさいので、 音が静かなモーターにしてほしい | ポジティブ ネガティブ ニュートラル |

評価用データと同一のドメインの投稿文を教師データに用いた。具体的には、あるドメインの投稿文 1,500 件中 1,200 件を学習に用い、残りの 300 件を評価用データとして用いた。

実験 2 同一ドメインの投稿文と不満投稿による学習

評価用データと同一のドメインの投稿文を教師データに用いた。また、教師データにはあるドメインの投稿文に加えて不満投稿を用いた。具体的には、あるドメインの投稿文と不満投稿各 600 件、計 1,200 件の投稿文を学習に用い、学習に用いなかったドメインの投稿文 900 件中 300 件を評価データとして用いた。

実験 3 異なるドメインの投稿文のみによる学習

評価用データとは異なるドメインの投稿文を教師データとして用いた。具体的には、あるドメインの投稿文 1,200 件を学習に用い、学習に用いたものとは異なるドメインの投稿文 300 件を評価用データとして用いた。

実験 4 不満投稿のみによる学習

教師データとして不満投稿のみを用いて各ドメインの投稿文の分類を行った。具体的には、不満投稿 1,200 件を学習に用い、各ドメインの投稿文 300 件を評価データとして用いた。

実験 5 異なるドメインの投稿文と不満投稿による学習

評価用データとは異なるドメインの投稿文を教師データとして用いた。また、教師データにはあるドメインの投稿文に加えて不満投稿を用いた。具体的には、あるドメインの投稿文と不満投稿各 600 件、計 1,200 件の投稿文を学習に用い、学習に用いたものとは異なるドメインの投稿文 300 件を評価用データとして用いた。

4.2.3 実験結果及び考察

Amazon の「日用品」、「家電機器」の投稿文を評価データとしてネガティブ感情の分類実験を行った結果を図 1, 図 2 にそれぞれ示す。また以下に、各実験の結果に対する考察を述べる。

実験 1 同一ドメインの投稿文のみによる学習

図 1 の青色のグラフである日用品、図 2 の青色のグラフである家電機器が実験 1 の結果である。結果より、実験 1 では全体的な精度こそ高いものの、日用品、家電機器どちらの場合においても再現率が適合率よりも低くなっている。これは、ネガティブにラベル付けされているデータのうち、“コストパフォーマンスが悪い”や“とても使いづらい”のような直接的なネガティブ表現は正しく分類できていたのに対し、“かなりの忍耐と、寛容な心が必要でしょう。”や“スプレーだけで落ちてほしかったなあ。”のような間接的なネガティブ表現が正しく分類できていなかった。さらに、ネガティブともニュートラルともとれるようなデータも正しく分類できていなかった。このため、再現率が低くなったと考えられる。実際に表 9 において、作業員間の一致度であるカッパ係数の値はどのデータにおいてもおよそ 0.4 から 0.5 の間の値を示している。こ

表 9: 感情付与結果

| データ | | | ポジティブ 節数 | ネガティブ 節数 | ニュートラル 節数 | カップ 係数 |
|--------------|------|---------------|---------------|---------------|----------------|-----------|
| サイト ドメイン | カテゴリ | 商品 | | | | |
| 不満買取 センター | 不満 | 不満投稿 | 69 (2%) | 2257 (50%) | 2152 (48%) | 0.435 |
| Amazon | 日用品 | トイレッ トペーパー | 955 (31%) | 442 (14%) | 1721 (55%) | 0.473 |
| | | 洗剤 | 1234 (28%) | 700 (16%) | 2492 (56%) | 0.504 |
| | | 防虫剤 | 777 (24%) | 370 (12%) | 2492 (64%) | 0.460 |
| | | 合計 | 2966 (27%) | 1512 (14%) | 6705 (60%) | 0.479 |
| Amazon | 家電機器 | ノート PC | 1054 (21%) | 858 (17%) | 3187 (62%) | 0.464 |
| | | ロボット 掃除機 | 1922 (25%) | 1051 (14%) | 4628 (61%) | 0.479 |
| | | スマート フォン | 1118 (20%) | 1086 (19%) | 3505 (61%) | 0.513 |
| | | 合計 | 4094 (22%) | 2995 (16%) | 11320 (61%) | 0.485 |

れは Landis ら [25] の示した指標によると中程度の一致 (Moderate) であり, 作業員間の感情評価の一致度が十分高いとはいえない. このことから, 今回用いたデータの中に作業員によって感情の評価が分かれるようなデータが一定数存在していることがわかる.

実験 2 同一ドメインの投稿文と不満投稿による学習

実験 2 では, 教師データにネガティブなデータの集合である不満投稿を加えることで, 同一ドメインにおけるネガティブ感情の分類精度にどのような影響が現れるかを確認する.

図 1 の緑色のグラフである日用品+不満投稿, 図 2 の緑色のグラフである家電機器+不満投稿が実験 2 の結果である. 結果より, 日用品, 家電機器どちらの場合においても, 不満投稿を加えなかった実験 1 の結果と比べて低い値を示している. 理由として, 不満投稿を教師データに加えることで, 実験 1 の同一ドメインのみでは正しく分類できていなかった“面倒”といったネガティブ表現を正しくネガティブと分類できるようにはなっていた. しかしながら, 不満投稿では“～が嫌”や“～が悪い”のような直接的なネガティブ表現が多いのに対し, Amazon の投稿文には“5 分待つ必要がある”, “薬品の匂いがする”といった間接的にネガティブを表しているようなデータも多く存在している. このため, 実験 1 と比べて精度が低くなったと考える.

実験 3 異なるドメインの投稿文のみによる学習

実験 3 では異なるドメインのデータを学習に用いた際, どの程度の精度が得られるかの確認を行う.

図1の赤色のグラフである家電機器，図2の赤色のグラフである日用品が実験3の結果である。結果より，日用品，家電機器どちらにおいても実験1で行った同一ドメインの投稿文による学習，評価よりも低い精度を示している。特に再現率に関しては今回行った5つの実験の中で最も低い。これはこの2つのドメインそれぞれのネガティブなデータにおいて，“ストレス”や“無駄”など両者に共通して出現しているネガティブな表現がある。その一方で日用品では“臭い”や“汚れ”，家電機器では“エラー”や“フリーズ”などといった，あるドメインに特有のネガティブな表現も存在している。これら共通ではないネガティブな表現のデータを正しく分類できていないために再現率が低くなったと考える。また，特に日用品のデータを学習に用いて家電機器のデータを分類した結果が，家電機器のデータを学習に用いて日用品のデータを分類した場合と比べて著しく低くなっている。これは，日用品のデータと比べて家電機器のデータには“ブルースクリーン”や“バグ”などのドメイン特有のネガティブな表現が多数存在しているためである。

実験4 不満投稿のみによる学習

実験4ではジャンルを問わないネガティブなデータの集合である不満投稿のみを学習に用いた際の分類精度を確認する。

図1と図2の黒色のグラフである不満投稿が実験4の結果である。今回学習に用いた不満投稿のネガティブなデータは“匂いがキツすぎてイヤ”など直接的な表現で書かれていることが多い。そのため，“虫が大量発生”(=気持ち悪い)や“ファンの音が大きい”(=うるさい)などの間接的な表現をうまく分類できなかった。その為，実験1や実験2と比べて再現率，適合率ともに低くなっている。

一方で，ドメイン固有のネガティブ表現に対応できなかった実験3と比べると，再現率は日用品，家電機器のどちらでも高い値を示している。これは不満投稿には日常における不満がジャンルを問わずに投稿されている。そのため，ドメイン毎に固有のネガティブ表現にもある程度対応ができていたことが要因だと考える。

実験5 異なるドメインの投稿文と不満投稿による学習

実験5では評価データと異なるドメインのデータ及び不満投稿を学習に用いることで，異なるドメインのデータに対するネガティブ感情分類精度，即ち学習器の汎用性がどの程度得られるかを確認する。

図1の黄色のグラフである家電機器+不満投稿，図2の黄色のグラフである日用品+不満投稿が実験5の結果である。日用品の投稿文に対する実験結果である図1の黄色のグラフを見ると，異なるドメインのみを用いた実験3の赤色のグラフや不満投稿のみを用いた実験4の黒色のグラフと比較すると再現率の値が上昇しているのがわかる。これは学習時に不満投稿を用いることで，異なるドメインのみでは対応できなかったネガティブ表現対応ができるようになったためと考える。さらに，異なるドメインのデータを用いることで不満投稿のみでは対応できなかった間接的なネガティブ表現に対応ができるようになったためとも考える。

一方で，図2の黄色のグラフを見ると，実験3と同様に再現率が大きく落ち込んでしまっている。日用品と不満投稿を学習に用いた場合では，不満と家電機器を用いた場合と異なり，間接的な表現に対してもドメイン固有の表現に対してもあまりうまく分類できなかった。また，不満投稿のみを用いている黒色のグラフである実験4では図1と図2で再現率に大きな差は見られない。これらを鑑みるに，実験5での学習において，日用品の投稿文と不満投稿の混合の場合では，うまく学習が行われず，そのため精度が下がってしまったと考える。これに対してはパラメータの調整を行うことで多少改善が図れるのではないかと考えられる。

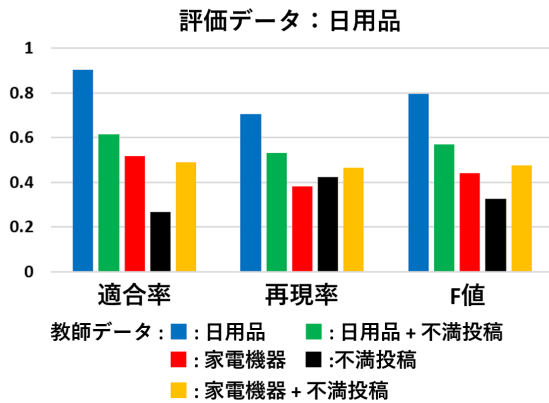


図 1: Amazon - 日用品 実験結果

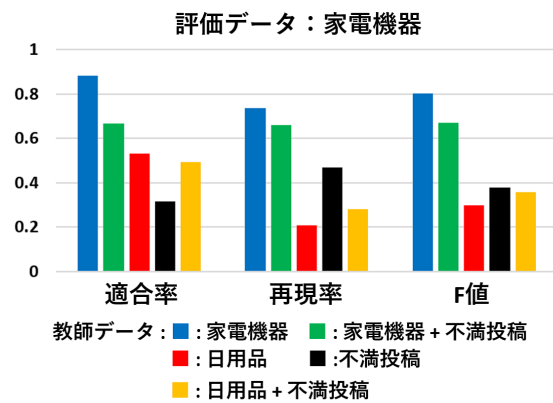


図 2: Amazon - 家電機器 実験結果

総評

以上5つの実験の結果より、学習と評価に同じドメインの投稿文を用いている実験1と異なるドメインの投稿文を用いている実験3の結果を比較すると、適合率、再現率ともに減少していることがわかった。一方で、実験3と実験5を比較すると、実験5の再現率のほうが実験3よりも高い値を示している。これらの結果より、今回学習に用いた不満投稿のデータがネガティブの感情分類において有用であることがわかる。しかしながら、実験1と実験5を比較すると、実験5の再現率は低くなっている。理由としてはドメイン固有のネガティブな表現や、直接的、間接的なネガティブ表現の表記の違いに完全に対応しきれていないことが考えられる。そのため、パラメータの調整や学習に用いる最適なドメインデータの選定、学習に用いるデータ毎の最適な割合などを調べてゆく必要があると考え、今後の課題である。

5 レビューの感情に基づいた評価の高いレビューの分析

ネットショッピングにおいて、ユーザはレビューを参考にして商品を購入する 경우가多数ある。ユーザは商品について書かれたレビューを読むことでその商品についてより深く理解し、購入するかどうかを判断する根拠のひとつとしている。また各レビューには他のユーザがそのレビューが参考になったかどうかの点数が表示されており、この点数から参考に成り得るレビューの目星をつけることがある。しかしながら、必ずしも全てのレビューに点数がついて評価されているわけではなく、参考になる内容であるにも関わらず埋もれてしまっているレビューも多数存在している。一方で安藤ら [5] はレビューに対して 3 つの大項目及び 23 の詳細項目からなる分類を行い、各項目が消費者の購買意欲に与える影響を評価している。この研究によると、商品レビューに関して商品に対する購入者の気持ち及び商品に対する客観的事実が他の項目と比べて影響を与えることが判明している。本研究では、この安藤らの研究を参考にし、消費者の購買意欲に影響を与えるレビューは評価の高いレビューであると考え、評価の高いレビューとレビュー中の購入者の気持ち及び商品に対する客観的事実に着目した分析を行う。尚、購入者の気持ち及び商品に対する客観的事実においてレビュー中で商品に対する購入者の気持ちが書かれている箇所はポジティブやネガティブの感情がよく表れていると考えられる。また一方で商品に対する客観的事実を述べている箇所では、商品の大きさや数量、スペックなどが主であり、ポジティブやネガティブではなくニュートラルの感情が多く存在していると考えられる。以上のことから本論文では、購入者の気持ち及び商品に対する客観的事実をその商品に対するレビュー文中のポジティブ、ネガティブ、ニュートラルの各感情の分布とみなし、レビュー中の感情の分布と評価の高いレビューとの関係性の分析を行っていく。またレビューには「この掃除機、性能は良いんだけど音がうるさい」という文中での「性能は良い」、「音がうるさい」のように 1 つの文中に複数の感情が混在しているものも多く存在している。そのため文章や文といった単位では正しくレビュー中の感情を捉えられないと考え、本分析では文を更に細かくした節という単位を 1 つの感情の範囲とする。

5.1 レビューの感情に基づいた評価の高いレビューの分析

本論文では、他者から参考にされている評価の高いレビューはレビュー中の感情の分布が影響していると考え、評価の高いレビューにおける感情に着目した分析を行う。本分析では 4 章で用いたレビューデータに加えて、他のユーザがレビュー内容を参考にすることが多いと考えられる化粧品カテゴリのレビューデータを新たに用いる。

5.2 データセット

分析には 4.2 章と同様に我々の行ったアンケートにより節毎にポジティブ、ネガティブ、ニュートラルいずれかの感情が付与されたレビューを用いる。今回は新たに化粧品カテゴリのレビューとして「化粧下地」のレビューを 500 件使用する。化粧品カテゴリのレビューの感情付与の方法は 4.2 章と同様に 10 人の作業員による感情の評価であり、過半数の評価が集まった感情を該当する節の感情とする。分析に用いるデータセットの例を表 10 に示す。

表 10: 分析に用いるデータセット

| データ | | | ポジティブ 節数 | ネガティブ 節数 | ニュートラル 節数 | カップ 係数 |
|-------------|------|---------------|---------------|---------------|---------------|-----------|
| サイト ドメイン | カテゴリ | 商品 | | | | |
| Amazon | 日用品 | トイレット ペーパー | 955 (31%) | 442 (14%) | 1721 (55%) | 0.473 |
| | | 防虫剤 | 777 (24%) | 370 (12%) | 2492 (64%) | 0.460 |
| Amazon | 家電機器 | ノート PC | 1054 (21%) | 858 (17%) | 3187 (62%) | 0.464 |
| | | スマート フォン | 1118 (20%) | 1086 (19%) | 3505 (61%) | 0.513 |
| Amazon | 化粧品 | 化粧下地 | 1122 (34%) | 631 (19%) | 1582 (47%) | 0.529 |

5.3 仮説

分析をするにあたり、以下の4つの仮説を立てた。そして、実際に評価の高いレビューとそのレビュー中の感情分布を基に以下の仮説が正しいかどうかを検証する。尚、本分析において評価の高いレビューとは他のユーザ3人以上から参考になったと評価を受けているレビューとする。

仮説 1: レビュー中でポジティブ、ネガティブ両方の感情が記載されているレビューは評価が高いレビュー中の感情がポジティブもしくはネガティブのどちらかに偏っているレビューは、商品に対する良い点もしくは不満点のみが書かれていることから、商品に対しての投稿者個人の感情のバイアスがかかってしまっている可能性がある。一方でポジティブ、ネガティブ両方の内容が書かれているレビューは公平な視点に基づいていることが期待できるため、他者の参考になりやすく、評価も高くなると考えられることから仮説1を導出した。尚、本分析において”両方の感情が記載されている”とは1レビュー中にポジティブ、ネガティブの節がそれぞれ1つ以上存在しているものを指す。

仮説 2: レビュー中の商品に対する感情がニュートラルの箇所が多いレビューは評価が高いレビュー中で感情の表れていない箇所というのは商品に対する投稿者の評価ではなく、大きさや重さ、量など商品そのものについての情報が記述されていると考えられる。そのため、ニュートラルが多いレビューは他のユーザに商品に関する情報をよく伝えられると考えられることから仮説2を導出した。本分析においてはニュートラルの節が3つ以上連続しているかどうかをニュートラルが多いレビューの判断基準とする。

仮説 3: 異なる商品カテゴリのレビューであっても評価の高いレビューの傾向は同じであるレビューにおいては商品によってよく使われる単語や表現こそ異なるものの、レビューを見るユーザが求めている情報としては良い点や悪い点、使用感やスペック等々などの商品でも大きく変わるものではないと考えられる。そこで、商品が異なっても評価の高いレビューの傾向は変わらないという仮説3を導出した。

5.4 分析結果と考察

各商品レビューにおける他者から参考になると評価された回数の一覧を表 11 に示す。ノート PC やスマートフォン、化粧下地は良いと評価された数 (以下、評価数という) が 50 以上のものも存在しているのに対し、トイレトペーパー、防虫剤は評価数 20-30 のものが最大となっている。これはトイレトペーパーや防虫剤は日常的に使用するものであり、レビューの中身もリピーターによる似た内容のものが多いため、そこまでレビューが重要視されていないためと考えられる。一方でノート PC やスマートフォンは高価であり、また故障や不具合がつきものであるということ、化粧下地は実際に自分の体に使用するものであるということから、ユーザがレビューの内容を重要視しているためだと考えられる。

表 11: 他者から参考になると評価された回数の一覧

| 評価数 | 商品 | | | | |
|-------|----------|-----|--------|---------|------|
| | トイレトペーパー | 防虫剤 | ノート PC | スマートフォン | 化粧下地 |
| 0 | 213 | 213 | 61 | 134 | 154 |
| 1 | 125 | 92 | 79 | 65 | 122 |
| 2 | 58 | 63 | 61 | 42 | 54 |
| 3 | 27 | 33 | 45 | 51 | 49 |
| 4 | 23 | 30 | 38 | 35 | 26 |
| 5 | 13 | 13 | 29 | 24 | 23 |
| 6-10 | 34 | 41 | 73 | 71 | 56 |
| 11-20 | 7 | 13 | 57 | 43 | 23 |
| 21-30 | 0 | 2 | 27 | 17 | 8 |
| 31-40 | 0 | 0 | 11 | 11 | 1 |
| 41-50 | 0 | 0 | 7 | 4 | 1 |
| 51- | 0 | 0 | 11 | 7 | 7 |

仮説 1: レビュー中でポジティブ、ネガティブ両方の感情が記載されているレビューは評価が高い各商品の誰からも参考にされていないレビューと評価の高いレビューにおける、ポジティブ、ネガティブ両方の感情が記載されている割合を表 12 に示す。また、レビュー中にポジティブの感情が記載されている割合とネガティブの感情が記載されている割合をそれぞれ表 13, 表 14 に示す。表 12 より、どの商品においてもポジティブ、ネガティブ両方の感情が記載されている割合は評価数が 0 の場合が最も低く、参考にしていく評価数が増えるに従って増加している。このことからポジティブ、ネガティブ両方の感情が記載されているレビューと他者から参考にされる数には関係性があることがわかる。一方で表 13, 表 14 を見ると、評価人数が増えるに従ってポジティブの感情の割合は減少しているのに対し、ネガティブの感情の割合は反対に増加している。このことから、レビューを見るユーザはネガティブな感情をより重視していることがわかる。よって仮説 1 は「真」であるが、追加情報として、個々の感情はユーザはネガティブな感情をより重視していることがわかった。

表 12: ポジティブ, ネガティブ両方の感情が記載されている割合

| 商品 | 評価人数 | | | | |
|-----------|-------|-------|-------|-------|-------|
| | 0人 | 3人以上 | 5人以上 | 7人以上 | 10人以上 |
| トイレットペーパー | 0.286 | 0.353 | 0.415 | 0.444 | 0.667 |
| 防虫剤 | 0.247 | 0.450 | 0.493 | 0.500 | 0.706 |
| ノート PC | 0.333 | 0.453 | 0.486 | 0.488 | 0.460 |
| スマートフォン | 0.391 | 0.525 | 0.564 | 0.598 | 0.679 |
| 化粧下地 | 0.411 | 0.537 | 0.573 | 0.643 | 0.667 |

表 13: ポジティブの感情が記載されている割合

| 商品 | 評価人数 | | | | |
|-----------|-------|-------|-------|-------|-------|
| | 0人 | 3人以上 | 5人以上 | 7人以上 | 10人以上 |
| トイレットペーパー | 0.408 | 0.323 | 0.309 | 0.223 | 0.252 |
| 防虫剤 | 0.368 | 0.245 | 0.199 | 0.198 | 0.195 |
| ノート PC | 0.291 | 0.219 | 0.176 | 0.148 | 0.139 |
| スマートフォン | 0.341 | 0.244 | 0.230 | 0.212 | 0.169 |
| 化粧下地 | 0.443 | 0.329 | 0.295 | 0.268 | 0.245 |

仮説 2: レビュー中の商品に対する感情がニュートラルの箇所が多いレビューは評価が高い
 各商品レビューにおける連続した3つ以上のニュートラルの節が存在している割合を表 15 に示す。こちらもどの商品においても参考になっている人数が増えるに従って割合が増加している。しかしながら、評価数が7人を超えると増加率は横ばいである。これは仮説 1 の表 12, 表 14 の割合が評価数が増えるに従い増加していることを併せて鑑みると、多くのユーザから参考にされているレビューでは、商品の解説等の客観的事実を表すニュートラルの感情だけでなく、商品の良い点や悪い点に関するポジティブ, ネガティブの感情も併せて記載されることが多く、相対的にニュートラルの割合が上がりづらくなっていると考えられる。このことからレビュー中に連続したニュートラル節の箇所が存在するレビューは評価が高くなりやすいことがわかる。しかし、より高い評価を得るためには客観的事実などのニュートラルだけでなく、ポジティブや

表 14: ネガティブの感情が記載されている割合

| 商品 | 評価人数 | | | | |
|-----------|-------|-------|-------|-------|-------|
| | 0人 | 3人以上 | 5人以上 | 7人以上 | 10人以上 |
| トイレットペーパー | 0.107 | 0.177 | 0.171 | 0.155 | 0.174 |
| 防虫剤 | 0.091 | 0.151 | 0.189 | 0.185 | 0.126 |
| ノート PC | 0.155 | 0.221 | 0.268 | 0.287 | 0.305 |
| スマートフォン | 0.198 | 0.204 | 0.215 | 0.218 | 0.246 |
| 化粧下地 | 0.146 | 0.227 | 0.245 | 0.270 | 0.311 |

ネガティブの感情も必要であることが同時にわかる。よって仮説2も「真」であるが、追加情報として、評価が高くなるにつれてニュートラルだけではなく、ポジティブやネガティブの感情も重要となってくるということがわかった。

表 15: 連続した3つ以上のニュートラルの節が存在している割合

| 商品 | 評価人数 | | | | |
|-----------|-------|-------|-------|-------|-------|
| | 0人 | 3人以上 | 5人以上 | 7人以上 | 10人以上 |
| トイレットペーパー | 0.359 | 0.406 | 0.538 | 0.559 | 0.556 |
| 防虫剤 | 0.388 | 0.674 | 0.701 | 0.675 | 0.765 |
| ノート PC | 0.434 | 0.595 | 0.616 | 0.650 | 0.645 |
| スマートフォン | 0.422 | 0.615 | 0.634 | 0.674 | 0.714 |
| 化粧下地 | 0.238 | 0.390 | 0.418 | 0.443 | 0.431 |

仮説3:異なる商品カテゴリーのレビューであっても評価の高いレビューの傾向は同じである

仮説3では商品の違いはレビューの評価の傾向に影響しないとした。しかし、仮説1の表12を見ると、ノートPCの結果は他の商品と違い大きな変化は見られない。これはノートPCのレビュー中にはパソコンに詳しいユーザが投稿したものとあまり詳しくないユーザが投稿したものが混在している。パソコンに詳しいユーザが投稿したレビューには「液晶は色は全く当てにできませんが、この手の液晶は高い色域100%物より、文書等見やすいので、長時間でも困ることは無いでしょう。」のようにポジティブ、ネガティブ両方の感情が記載されていることが多い。一方で、詳しくないユーザが投稿したレビューは「購入してから3ヶ月、問題なく動いているので満足です」のように片方の感情のみが記載されていることが多い。しかし、パソコンに詳しくないユーザにとってはこのような情報も有益であるため、詳しくないユーザのレビューにも多くの評価がされていた。このため評価数と感情との間にはっきりとした関係性が見られなかったと考えられる。また、表11においても商品のジャンルによって評価数の分布が異なっていることがわかる。このように商品によってレビューを参考にするユーザの層が異なるため、商品の違いが高評価のレビューの傾向に影響している。よって仮説3は「偽」である。

上記のように、レビューにおいては複数の感情が混在している方が評価が高くなる傾向にあることがわかった。また、評価が高くなるにつれてレビュー中のポジティブの割合は減少するのに対しネガティブの割合は増加していることから、ネガティブの感情がより重要視されていることがわかる。さらに、商品ごとの評価数の分布などを見てもわかるように、今回商品によって評価の傾向が異なることが判明したため、今後は商品の違いも考慮していく必要があると考えられる。

6 まとめと今後の課題

6.1 機械学習を用いた感情抽出に基づく潜在的フォロワー抽出手法

本研究ではよりよいフォロワーを推薦するためにツイート内容の類似に加え、機械学習を用いて類似する内容に対する感情の類似も考慮した潜在的フォロワー推薦手法を提案し、従来手法との比較実験を行うことで提案手法の有用性を示した。また、提案手法で潜在的フォロワーを推薦できることを確認した。その結果、提案手法のほうが従来手法よりも良い精度であったが、まだ十分でなく、改善していく必要があるといえる。

今後の課題としては、SVMの教師データ数を増やす、ツイートが複数の感情を持つ場合や感情を持たない場合などにも対応ができるようにする、ユーザのTwitterの利用傾向も推薦時に考慮する、文章だけでなく画像なども分析を行っていくなどがある。

6.2 レビューにおけるCRFを用いた節単位毎の感情分類手法

次に、学習の際に、あるドメインの投稿文に新たにネガティブなデータを加えることで、異なるドメインのネガティブデータ分類において汎用性が得られ、教師データの作成コストの削減ができるかの確認を行った。実験にはAmazonの商品レビューと不満買取センターに投稿された不満投稿を用いた。結果としては、教師データにネガティブデータを補完することでネガティブデータの分類が向上するケースを確認できた。しかしながら、同一のドメインによる学習、分類の精度を超えることはできなかった。理由としては学習に用いるデータの選定およびパラメータの調整が不足していたことが挙げられる。

そのため今後の課題として、マルチドメインの分類に適しているデータの調査やパラメータ調整、そしてアンサンブル学習など他の手法との比較などを行っていききたい。

6.3 レビューの感情に基づいた評価の高いレビューの分析

そして、レビュー文中の感情の分布と他者からの評価数との関係性についての分析も行った。結果としてレビュー文中にポジティブ、ネガティブ両方の感情が存在してる場合や、ニュートラルの箇所が多いレビューは評価が高くなりやすい傾向にあることが確認できた。また、ネガティブの感情は評価が高くなるにつれて重要性が上がるということもわかった。

そして商品の違いが評価の傾向に影響を及ぼすことが判明したため、今後は商品の購買層や性質などの違いも考慮した分析を行っていききたい。

謝辞

本論文は，平成 29 年 4 月から平成 31 年 3 月までの間，甲南大学自然科学研究科修士課程知能情報学専攻に在学中，同専攻灘本研究室で行った研究の成果をまとめたものである。

本研究を進めるにあたり，大変多くの方にお世話になりました。ここに深く感謝を表します。研究に際して，約 3 年半もの期間にわたりご指導をいただきました恩師，灘本明代先生に心より深く感謝申し上げます。そして論文にまとめるにあたり，有益な御助言とご教示を賜りました甲南大学 若谷彰良先生，甲南大学 新田直也先生に心より謝意を申し上げます。また，日頃の研究会において多くのご指摘を下さいました灘本研究室の同期の皆様，実験の際に被験者を快く引き受けてくださった後輩の皆様に深く感謝いたします。さらに，千葉工業大学 熊本忠彦先生，株式会社 Insight Tech 三澤賢祐様，株式会社 Insight Tech 成田和弥様には研究を通して大変お世話になり，また貴重なアドバイスも数多くいただきましたこと，感謝申し上げます。最後に，大学院に進学するという決断に快く背中を押して下さり，ありとあらゆる場面で私を温かく見守り，続けてくれた父 秋山郁雄，母 秋山麻子に深く感謝致します。本当にありがとうございました。

研究業績

国際会議

- Kazuhiro Akiyama, Tadahiko Kumamoto, Akiyo Nadamoto, “ Emotion-based Method for Latent Followee Recommendation in Twitter, ” The 19th International Conference on Information Integration and Web-based Applications & Services(iiWAS2017), pp. 121-125, 2017.
- Kazuhiro Akiyama, Kensuke Mitsuzawa, Kazuya Narita, Tadahiko Kumamoto, and Akiyo Nadamoto, “ Clause-level Negative-opinion Analysis for Classifying Reviews on Multiple Domains, ” The 20th International Conference on Information Integration and Web-based Applications & Services(iiWAS2018), pp. 113-121, 2018.

国内会議

- 秋山和寛, 熊本忠彦, 灘本明代, “ ツイートからの多次元感情抽出手法の考察, ” 第8回ソーシャルコンピューティングシンポジウム (SoC2017), 信学技報, Vol. 117, No. 108, DE2017-3, pp. 11-16, 2017.
- 秋山和寛, 熊本忠彦, 灘本明代, “ 機械学習を用いた潜在的フォロイー抽出手法の提案, ” 第10回 Web とデータベースに関するフォーラム (WebDB Forum 2017), 信学技報, Vol. 117, No. 212, DE2017-16, pp. 51-56, 2017.
- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “ CRF を用いたレビューにおける節単位毎の感情推定, ” 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), J2-4, 7 pages, 2018.
- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “ 複数ドメインの消費者投稿文の節単位感情分析における手法の比較検討, ” 第9回ソーシャルコンピューティングシンポジウム (SoC2018), 信学技報, Vol. 118, No. 107, DE2018-4, pp. 15-20, 2018.
- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “ CRF を用いた複数ドメインの消費者投稿文におけるネガティブ感情分類, ” 第11回 Web とデータベースに関するフォーラム (WebDB Forum 2018), 信学技報, vol. 118, no. 213, DE2018-18, pp. 55-60, 2018.
- 秋山和寛, 熊本忠彦, 灘本明代, “ レビューの印象に基づいた評価の高いレビューの分析, ” 第11回データ工学と情報マネジメントに関するフォーラム (DEIM2019), 2019(to appear)

ポスター発表

- 秋山和寛, 熊本忠彦, 灘本明代, “ 機械学習を用いた潜在的フォロイー抽出手法の提案, ” 第10回 Web とデータベースに関するフォーラム (WebDB Forum 2017), 信学技報, Vol. 117, No. 212, DE2017-16, pp. 51-56, 2017.

- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “CRF を用いたレビューにおける節単位毎の感情推定,” 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), J2-4, 7 pages, 2018.
- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “CRF を用いた複数ドメインの消費者投稿文におけるネガティブ感情分類,” 第11回 Web とデータベースに関するフォーラム (WebDB Forum 2018), 信学技報, vol. 118, no. 213, DE2018-18, pp. 55–60, 2018.

研究業績-表彰

学生プレゼンテーション賞

- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “CRF を用いたレビューにおける節単位毎の感情推定,” 第10回データ工学と情報マネジメントに関するフォーラム (DEIM2018), J2-4, 7 pages, 2018.

学生奨励賞

- 秋山和寛, 三澤賢祐, 成田和弥, 熊本忠彦, 灘本明代, “CRF を用いた複数ドメインの消費者投稿文におけるネガティブ感情分類,” 第11回 Web とデータベースに関するフォーラム (WebDB Forum 2018), 信学技報, vol. 118, no. 213, DE2018-18, pp. 55–60, 2018.

参考文献

- [1] Scott A. Golder, Sarita Yardi, Alice Marwick, and Danah Boyd, “A Structural approach to contact recommendations in online social networks”, Proceedings of Workshop on Search in Social Media at ACM SIGIR Conference on Information Retrieval, 2009.
- [2] 大村涼, 赤石美奈, 佐藤健, “語彙構造を用いた Twitter ユーザ推薦手法の提案”, 情報処理学科全国大会講演論文集 2013(1), pp.609–611, 2013.
- [3] Jilin Chen, Werner Geyer, Casey Dugan, Michael Muller and Ido Guy, “Make new friends, but keep the old: recommending people on social networking sites”, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.201–210, 2009.
- [4] 久米雄介, 打矢隆弘, 内匠逸, “興味領域を考慮した Twitter ユーザ推薦手法の提案と評価”, 情報処理学会研究報告 (知能と複雑系), 2015-ICS-179 (1), pp.1–8, 2015.
- [5] 安藤まや, 関根聡, “レビューには何が書かれていて、読み手は何を読んでいるのか?”, 言語処理学会 第 20 回年次大会 発表論文集, pp.884–887, 2014.
- [6] Marcelo G. Armentano, Daniela L. Godoy, Analia A. Amandi, “A topology-based approach for followees recommendation in twitter”, A structural Approach to Contact Recommendations in Online Social Networks, 2011.
- [7] 北村太一, 小川祐樹, 諏訪博彦, 太田敏澄, “コミュニケーションに着目した Twitter フォロワーユーザ推薦”, 人工知能学会全国大会論文集 (CD-ROM), Vol. 26, 2012.
- [8] 黒柳智士, 山田泰宏, 鈴木浩, 服部哲, 速水治夫, “著名人情報に基づいた Twitter フォロワーユーザ推薦システム”, 研究報告グループウェアとネットワークサービス (GN) 2013-GN-86(25), pp.1–4, 2013.
- [9] 桑原雄, 稲垣陽一, 草野奉章, 中島伸介, 張建偉, “マイクロブログを対象としたユーザ特性分析に基づく類似ユーザの発見および推薦方式”, 情報処理学会研究報告 (データベースシステム研究会), Vol.149, No. 18, pp. 1–3, 2009.
- [10] Samir Rustamov, Elshan Mustafayev, Mark Clements, “Sentence-level subjectivity detection using neuro-fuzzy models”, Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pp.108–114, 2013.
- [11] Georgios Paltglou, Mike Thelwall, “More than bag-of-words:sentence-based document representation for sentiment analysis”, Proceedings of Recent Advances in Natural Language Processing, pp.546–552, 2013.
- [12] 田村一樹, 吉川大弘, 古橋武, “評点付きレビュー文書を対象としたトピックモデルの構築に関する検討”, 情報処理学会論文誌 Vol.56, No.3, pp.1013–1027, 2015.
- [13] 鈴木泰裕, 高村大也, 奥村学, “Weblog を対象とした評価表現抽出”, 人工知能学会セマンティックウェブとオントロジー研究会, 2004.

- [14] Chenghua Lin, Yulan He, “Joint sentiment/topic model for sentiment analysis”, Proceedings of the 18th ACM conference on Information and knowledge management, pp.375–384, 2009.
- [15] Yejin Choi, Claire Cardie, “Learning with compositional semantics as structural inference for subsentential sentiment analysis”, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pp.793–801, 2008.
- [16] Ying Zhao, George Karypis, “Comparison of agglomerative and partitional document clustering algorithms”, University of Minnesota, pp.2–14, 2002.
- [17] Mizuki Fujisawa, “Bayon - a simple and fast clustering tool - google project hosting”, 2012, <https://code.google.com/p/bayon/wiki/Tutorial/>, [Online; accessed 12-February-2019].
- [18] Yuki Yamamoto, Tadahiko Kumamoto and Akiyo Nadamoto, ”Multidimensional sentiment calculation method for Twitter based on emoticons” , International Journal of Pervasive Computing and Communications, Vol.11 Iss:2, pp.212–232, 2015.
- [19] Chih-Chung Chang, Chih-Jen Lin, “LIBSVM – A library for support vector machines”, 2014, <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>, [Online; accessed 12-February-2019].
- [20] 中村明, 感情表現辞典, 東京堂出版, 1993.
- [21] 熊本忠彦, 灘本明代, “共通話題に対する感情的態度の類似度に基づくフォロワー推薦”, 電子情報通信学会論文誌, Vol.J100-D, N0.4, pp.500–509, 2017.
- [22] 岡崎直観, CRFsuite: A fast implementation of Conditional Random Fields, 2007, <http://www.chokkan.org/software/crfsuite/>, [Online; accessed 12-February-2019].
- [23] 国立国語研究所 (2004), “分類語彙表増補改訂版データベース (ver.1.0) ”, 2015.
- [24] 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.
- [25] J Richard Landis, Gary G. Koch, “The measurement of observer agreement for categorical data.”, Biometrics, Vol.33 No.01 pp.159–174,1977.