

コメントフィルタリングのための感情分析を用いたコメント評価尺度の検討

笠原璃音¹ 大和淳司¹

¹工学院大学大学院工学研究科情報学専攻

em23013@ns.kogakuin.ac.jp, yamato@cc.kogakuin.ac.jp

概要

本研究では、インターネット上でのコメント炎上やエコーチェンバー現象に対処するため、閲覧者が自らコメントを調整できるシステムを目指し、コメントフィルタリングの検討を行った。具体的には、TF-IDF と極性辞書を用いてコメントを数値化し、これを適切・不適切なコメントを分けるための尺度とした。提案した尺度の有用性を検証し、コメントの適切・不適切性を決定するコメントの特徴量を分析した。

1 はじめに

インターネット上でニュースを読む際、ニュース記事に対するコメントの投稿や閲覧が一般的となってきた。コメント投稿は本名である必要はなく、その匿名性の高さから躊躇なく誹謗中傷コメントを残すユーザが多い。SNS や Web ニュースのコメント欄では、誹謗中傷やエコーチェンバー現象が頻発し、炎上は大きな社会問題となっている。炎上やエコーチェンバーの抑止のために、コメント投稿者側にリテラシーを求めるのはもちろん、閲覧者側も不快な思いをしないために炎上やエコーチェンバーを回避する必要があると考える。本研究では、閲覧者側が表示させたいコメントの適切度合いを調整し、自らコメントフィルタリングをするようなシステムを提案する。

具体的にはコメントを数値化することで、適切・不適切なコメントを評価する尺度を設定した。本研究では、適切なコメントを『ポジティブかつ中身のあるコメント』とし、『ポジティブかつ中身のある度合い』を数値化したものをスコアと呼ぶ。スコア算出には TF-IDF と極性辞書による感情分析を利用し、本論文ではスコアの有用性検証やコメントの適切性が影響する特徴量の分析、さらにストップワードの除外がコメントフィルタリングに与える影響を検証する。

2 先行研究

狩野ら[1]は、SNS のリテラシー教育において教員の負担軽減を図るため、不適切な Twitter の投稿を検出する手法を提案している。投稿文を単語に分割し、各単語に TF-IDF 値と感情極性値を与える。その後、TF-IDF 値が閾値以下の単語と感情極性値が 0 の単語を除外し、残った単語の感情極性値の平均を投稿文の評価値としている。この手法は、重要度が高い単語の極性値のみを扱っているが、本研究では単語の TF-IDF 値と極性値の両方を考慮して評価値を算出する。また、先行研究を参考に、文章中の重要でない単語を除外し、後述する提案手法 2 として評価値を算出する。

竹内ら[2]の研究は、ネットいじめ防止を目的にニコニコ動画の誹謗中傷コメントの発見に焦点を当て、「炎上動画判別機」を開発し炎上動画を自動検出した。竹内らは、荒れている動画と荒れていない動画それぞれから約 800 件ずつのコメントを収集し、これらのコメントに感情分析 API を使用して感情分析を行った。得られた数値を動画のベクトルとし、SVM や決定木等を用いて学習モデルを構築し、炎上動画分類器を作成している。本研究では、教師データとして炎上コメントを予め収集せずに、コメントの評価値を算出することで、適切・不適切コメントの判定方法を検討する。

佐藤ら[3]が開発したシステムは、初等教育のグループワークをサポートするもので、音声認識によって対面の話し合いを記録し、GPT-3 を使用して残すべき発言を選別している。GPT-3 の学習データには、表示すべき発言、不要な発言、曖昧な発言の 3 つのクラスにラベル付けされたデータを再学習させたものを用い、発言内容を 3 クラス分類している。結果として、不要な文として選ばれたテキストは相槌が中心で、佐藤らの予想通りの動作を示した。GPT シリーズは特別な知識が不要で、多様なジャンルの文章に対し文脈に基づいた高精度なテキスト分類が可

能であるが、分類した結果のみが表示され、分類の根拠が不明瞭である。そのため個人によっては結果に納得できるが、一部の人には理解が難しい。本研究では、適切・不適切の二値評価に加えて、スコアとしてコメントの適切度合を示すことで、分類結果の根拠を説明する要因を導入している。

3 方法

提案するスコア算出手順を図1に示す。

3.1 コメント中の単語の極性を判定する

極性辞書を用いて単語のネガティブ・ポジティブを判定する。本研究で用いた極性辞書は乾ら[4]の日本語評価極性辞書である。

MeCabで形態素解析を行い、コメント中の名詞・動詞・形容詞のみを抽出し単語の極性判定を行う。図2のように、ネガティブな単語はn、ポジティブは単語がp、極性を持たない単語は-と出力され、この判定結果をもとに、後述する重みづけ計算のために単語ごとに係数を定める。今回は、ネガティブな単語は-1.1、ポジティブは単語が1.1、極性を持たない単語1.0と係数を設定した。

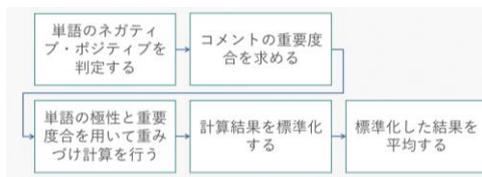


図1.スコアの算出手順

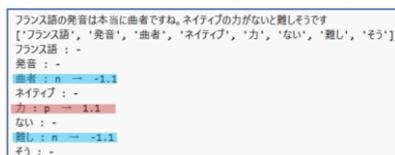


図2.コメントの形態素解析と極性判定結果

3.2 コメントの重要度を TF-IDF で求める

TF-IDF(Term frequency-Inverse document frequency)とは、文書中に含まれる各単語が、その文書内でどれくらい重要かを表す尺度である。TFとは『あるコメントの集まりにおける、ある単語の出現頻度』を表し、IDFとは『ある単語を含むコメントがコメントの集まりの中にどれくらい少ない頻度で存在するか』を表し、TFとIDFの乗算によって求められる。それぞれの求め方は式(1)の通りである。

$$TF = \text{コメント群における、ある単語の出現頻度}$$

$$IDF = \log\left(\frac{\text{総コメント数}}{\text{ある単語を含むコメントの数}}\right) + 1 \quad \dots(1)$$

ある単語が複数のコメントで使われていたらTF-IDFは低くなり、特定のコメントの中でのみ使われ、他のコメントでは使われていなかったらTF-IDFは高くなる。図3のように、文章を構成する一つの単語に対し一つのTF-IDFがあたえられる。TF-IDFは本来、文書内のある文章の重要度の算出や、文書の類似度の算出に用いられるが、今回は“中身のあるコメント=コメントの重要度”が高いと仮定し、TF-IDFを“中身のあるコメント”の尺度として採用した。

3.3 スコアを算出する

図4のように、極性を持つ単語に対応したTF-IDF値に、3.1節で極性単語に設定した係数を重み付け計算し、単語のTF-IDFに極性情報を付与する。

そして、分散具合を考慮するために、重みづけ計算を行ったTF-IDF値を標準化する。標準化したTF-ID値を平均した結果がスコアである。一連の算出方法を式(2)に示す。

$$Score = \frac{1}{n} \sum_{i=1}^n z_i, \quad z_i = \frac{x_i - \mu}{\sigma} \quad \dots(2)$$

3.3.1 提案手法1

3.1節から3.3節で説明したスコアの算出方法を提案手法1とする。

3.3.2 提案手法2

3.1節で形態素解析を行う際、『です』『ます』『が』『の』などのストップワードや、コメント群で頻繁に出現する単語を除去し、採用された単語にのみ極性判定を行いTF-IDFを割り当てる。その後の算出方法は提案手法1と同様である。以降、提案手法2で算出されたスコアをストップワード除外スコアとし、以降“STPW除外スコア”と呼ぶ。

フランス語	の	発音	は	本当	に	曲者	です	ね	.
1.6094	1.6178	1.6094	1.2992	1.6094	0.95551	1.6094	2.725	1.6094	1.0986
ネイティブ	の	力	が	ない	と	難し	そう	です	
1.6094	1.6178	1.6094	1.0986	1.0986	1.2992	1.6094	1.6094	2.725	

図3.各単語にTF-IDF値の割り当て



図 4.重みづけ計算の例

4 データセット作成

4.1 コメントの収集

Yahoo!ニュースのコメント欄に投稿されたコメント 102 件を収集した。具体的なニュース記事のタイトルについては付録 1 に記載した。

4.2 人手によるコメント評価

被験者 5 名に 102 件のコメントを提示し、『適切か不適切か』を判定するアンケートを行った。アンケートの教示内容と質問例を付録 2 に示す。アンケートで適切・不適切だと判定されたコメント例を付録 3 表 A に示す。アンケートで適切に投票した人の割合をアンケート結果として扱い、適切判定の投票数と総投票数の比を用いた。

4.3 コメントの特徴量

本研究では、コメントの分析を行うためにコメントの特徴量に“ポジティブ率”，“いいね率”，“正規化スコア(正規化 STPW 除外スコア)”を選定した。

- ポジティブ率：あるコメントに含まれているポジティブな単語数とあるコメントの単語数の比である。
- いいね率：Yahoo!ニュースのコメント欄には、コメントにいいね・バッドを付ける機能がある。あるコメントに付けられたいいねの数と、いいね数とバッド数の和の比である。
- 正規化スコア/正規化 STPW 除外スコア：スコアは単語数の長短によって影響を受けやすい。単語数を考慮するため、あるコメントの単語数をコメント群の総単語数で正規化し、あるコメントのスコアを乗算した数値を正規化スコアとする。

5 実験

5.1 実験 1

スコアの有用性を確認するために、スコアとアンケート結果で二次元散布図を作成し、二次元散布図

の第 2，第 4 象限に分布しているコメントを分析する。

提案手法 1 で算出したスコアとアンケート結果、正規化スコアとアンケート結果で二次元散布図を作成する。第 2，第 4 象限に分布しているコメントは不適切であると仮定し、第 2，第 4 象限に分布するコメントを抽出する。各特徴量の平均(Comment feature average:CFA とする)を求め、平均より下回っていたコメントの割合(Percentage below the comment feature average:PBCFA とする)を調べる。

提案手法 2 においても同様に、スコアと正規化スコアを STPW 除外スコア、正規化 STPW 除外スコアに変え、コメントを分析する。

5.2 実験 2

コメントの特徴量がコメントの適切・不適切性に影響しているかどうかの検証を行うために、ランダムフォレストを用いて、特徴量をもとにアンケート結果の予測を行う。

アンケート結果を目的変数、ポジティブ率・いいね率・スコア・正規化スコアを説明変数とし、ランダムフォレストを行い、アンケート結果を予測した。同様に、スコアと正規化スコアを STPW 除外スコア、正規化 STPW 除外スコアに変え、アンケート結果を予測した。

102 件のコメントの内、学習データを 8 割、テストデータを 2 割に分割し、5 回のクロスバリデーションを実施した。最終的な予測正答率は 5 回の平均値を使用している。

6 結果

6.1 実験 1

CFA と、スコアとアンケート結果の二次元散布図で第 2，第 4 象限に分布していたコメントの PBCFA を表 1 の“PBCFA...(A)”に示す。正規化スコアとアンケート結果での PBCFA を表 1 の“PBCFA...(B)”に示す。STPW 除外スコアとアンケート結果での PBCFA を表 1 の“PBCFA...(C)”に示す。正規化 STPW 除外スコアとアンケート結果での PBCFA を表 1 の“PBCFA...(D)”に示す。

表 1(A)において、CFA を下回ったコメント特徴量が、ポジティブ率と正規化スコア、アンケート結果の 3 件。表 1(B)においてポジティブ率と正規化スコアの 2 件となった。表 1(C)において、特徴量平均を

下回ったコメントの特微量がいいね率以外の4件。表1(D)においてはポジティブ率と正規化STPW除外スコアの2件となり、CFAを下回ったコメントが多かった。提案手法1, 提案手法2いずれにおいても、正規化をするとPBCFAが下がった。

6.2 実験2

提案手法1を用いてアンケート結果の予測を行ったときの正答率と、提案手法2を用いてアンケート結果の予測を行ったときの正答率を表2に示す。

表2. アンケート結果予測の正答率

	提案手法1	提案手法2
positive rate	0.69	0.69
like rate	0.46	0.46
Score	0.54	0.62
Score_norm	0.54	0.62
P_r · L_r · S	0.54	0.69
P_r · L_r · S_norm	0.62	0.54

提案手法1, 2において、もっとも正答率が高かったのは、説明変数にポジティブ率のみを用いたときだった。同率で、説明変数にポジティブ率, いいね率, STPW除外スコアを用いたときだった。

提案手法2においては、説明変数に特微量とSTPW除外スコアを用いたときは正答率が高くなるが、正規化STPW除外スコアを用いたときは正答率が下がることが分かった。

一方で、提案手法1においては説明変数に特微量と正規化スコアを用いると正答率が上がることが分かった。

7 考察

7.1 実験1

表1, 2より、スコアと正規化スコアのどちらにおいても、特微量の平均を下回ったコメントの割合が少なかったことが確認されたため、提案手法1はコ

メントの不適切性を判断する材料にはならないことが考えられる。

一方、STPW除外スコアにおいて、提案手法1よりも特微量の平均を下回ったコメントの割合が多かったため、提案手法2はコメントの不適切性を判断する材料になることが考えられる。

7.2 実験2

表2より、提案手法1, 2の両方においてポジティブ率がコメントの適切性に大きく関わり、いいね率はあまり関わらないことが考えられる。説明関数にいいね率のみを用いたときよりも、スコアもしくは正規化スコアを用いたときの方が正答率が上がったことから、スコアはコメントの適切性を説明する際に有用であることも考えられる。

提案手法1では、説明関数に特微量と併せてスコアの正規化を用いると正答率が上がったため、提案手法2でも同様の傾向がみられると予測していたが、提案手法2では正答率が下がった。この理由については不明で現在分析中である。

8 まとめ

本研究では、コメント閲覧者がフィルタリング強度を設定できるようなシステムを目指し、コメントの適切・不適切性を明確にするために、TF-IDFと感情分析を用いたコメントの数値化を行った。本論文では、コメントの尺度としてSTPW除外スコアが有用である可能性が確認された。また、コメント特微量を併せることで、フィルタリング精度は上がることが確認されたが、STPW除外スコアにおいては精度が下がることが確認された。提案手法によってフィルタリング結果の根拠を説明できる可能性が確認されたので、細かな尺度調整によるコメントの抽出が可能になると考える。また、将来的には複数感情を組み合わせた詳細な感情分析を行い、皮肉めいたコメントに対して文脈を考慮した分析を検討している。

表1. 第2, 4象限に分布するコメントのCFAとPBCFA

	positive rate	like rate	Score	Score_norm	Q_results
CFA(手法1)	0.06	0.78	1.44	14.50	0.84
PBCFA…(A)	0.56	0.19	0.44	0.59	0.93
PBCFA…(B)	0.50	0.36	0.39	0.68	0.31
CFA(手法2)	0.03	0.78	1.37	14.15	0.84
PBCFA…(C)	0.60	0.33	0.67	0.67	0.56
PBCFA…(D)	0.59	0.36	0.45	0.71	0.35

参考文献

1. 狩野, 柏熊, 山口, 河野, マッキン” ソーシャルメディアにおける感情極性を用いた文章の適正判定” 第 28 回ファジィシステムシンポジウム 2012 年 9 月
2. 竹内, 伊東 “文書分類手法による炎上動画検出手法の検討” 情報処理学会研究報告 2021 年
3. 佐藤, 白松, 吉村, 大森, 水本” GPT-3 を用いた音声認識結果の自動選別機構の小学生の対面議論への応用可能性” 人工知能学会第二種研究会資料 2022 年
4. 東北大学 乾・岡崎研究室 “日本語評価極性辞書”
(https://www.cl.ecei.tohoku.ac.jp/Open_Resources-Japanese_Sentiment_Polarity_Dictionary.html)

付録

付録 1

コメントを収集した Yahoo!ニュース記事のタイトル：

- 「初代ポケモン」を買いに!_外国人観光客が増加 何しに日本に【N スタ解説】(TBS NEWS DIG Powered by JNN)
- これが「タワマン」の未来か…かつては憧れだった「団地型マンション」のヤバすぎる現状 (現代ビジネス)
- 清原氏 中日来季ヘッドに急浮上 貧打解消へ劇薬プラン 立浪監督の PL 大先輩にコーチ&選手との緩衝材の期待 (デイリースポーツ)

付録 2

質問紙に掲載した教示内容：

- 適切かそうでないかの判定は『肯定的な文章だから』『中身のあるコメントだから』『ポジティブな単語が使われているから』などを基準にし、私情を含まず判定すること。『〇〇が嫌いだから・好きだから』『〇〇を知らないから』という理由で判定しないこと。
- 全てのコメントは Yahoo!ニュースのコメント欄からそのまま引用しており、誤字脱字や文法的な間違いがある可能性があるため、文法的な不備は考慮せず、判定を行うこと。
- 第三者の視点に立って判定を行うこと。

秋葉のレトロゲー売ってるところが外国人相手の商売価格に設定してるね。相場 * そもそもタワマンに住む気がしないね。地震大国日本なのに。 * の2倍だから全然日本人客入って来ないな

OK

OK

NG

NG

図 A.質問の例

付録 3

表 A.適切・不適切だと判定されたコメントの例

適切	家のどこかに眠っているかも！！ ゲームボーイ&カセット。 高額で売れないかな(^-^)
	日本の文化に触れてもらえることは素晴らしいことだと思います。
	清原さんを応援しております。清原さんには野球が1番輝く場所だと思っております。身体と精神的なものも考慮しながら清原さんにも活躍できる場所があればと思います。無理だけはされなくてほしいです。清原さんのご活躍をこれからも応援させていただきます！
不適切	タワマン叩く記事はページビュー稼げるから次々でるよねー。でも団地と同レベルで語るのって無理がありすぎない？
	つうか、訪日する外国人が日本語喋れるようにしてこいよ。ニューヨークで「日本語喋れないのか？」って店員に聞くか？
	今は抜け出しているとしても、何年もの間覚醒剤を使っていた人間をコーチにしてもいいものなのだろうか。使えるお金ができたのなら、また、覚醒剤に手を出さないという保証は全くないと思う。覚醒剤に手を出したなら、一生涯その誘惑に耐えていかななくてはならないと言われているのに、覚醒剤を買えるお金を与えてはいけないと思う。野球界で覚醒剤を覚え、その野球界に戻してはならないとも思う。