

# 感情分析を用いた口コミの NG ワード検知システムの開発

柿田秋祐<sup>†1</sup> 山本 堅嗣宣<sup>†2</sup> 山本純子<sup>†2</sup> 大貫勉<sup>†2</sup> 床井真理子<sup>†2</sup>  
原紳<sup>†3</sup> 渡邊信一<sup>†3</sup>

**概要:** インターネット上の口コミを NG ワードの検知により分類する際、単語の単純検索では投稿者の意図しない使われ方をしているものまで対象になってしまう場合がある。本研究は感情分析と形態素解析を用いてより正確に分類することを目的とする。

## Banned Words Detection from Reviews using Sentiment Analysis

SHUSUKE KAKITA<sup>†1</sup> MITUNOBU YAMAMOTO<sup>†2</sup> JUNKO YAMAMOTO<sup>†2</sup>  
TSUTOMU OHNUKI<sup>†2</sup> MARIKO TOKOI<sup>†2</sup>  
SHIN HARA<sup>†3</sup> SHINICHI WATANABE<sup>†3</sup>

### 1. 緒 言

#### 1.1 研究の背景

昨今、SNS や口コミサイトでの利用者数は増加の一途を辿っている。日本国内における SNS の利用者数は、2012 年度の 4,965 万人から 2017 年は 7,204 万人と大きく増加している。また、ネット利用人口に対する SNS 利用率も 2012 年から 2017 年まで約 20% 上昇しており、多くの人がインターネット上の情報に触れることができる時代となった[1]。

その中、インターネット人口が増加したことによって悪質な書き込みも増加している。従来は目視で口コミを確認していたが、増加する一方の口コミをすべて確認するのは非常に時間がかかる。よって、自然言語処理を用いた口コミの自動分類の需要が高まっている。

#### 1.2 NG ワード検知とは

NG ワード検知とは、あらかじめ禁止語句を NG ワードとして登録し、口コミ文中にその単語が現れたらそれを検知する方法である。この方法はほとんどの投稿サイトの運営会社で採用されている。しかし、NG ワードのリスト・検索条件共に無関係の研究者が手に入れることは難しい。またそれに伴って関係論文も少ないのが現状である。

#### 1.3 従来のシステム

本論文では、口コミサイトの「栃ナビ!」を運営するヤマゼンコミュニケーションズ (株) との共同研究の一環として行なった。今まで使われていた NG ワード検知方法は、文中に NG ワードがそのまま含まれているかのみを判別するものだった。しかしこの方法では、投稿者が意図していない使われ方をしているものも含まれてしまうことがある。例えば「無料」という意味の「ただ」(名詞) という単語が NG ワードに含まれている。従来は「いただきます」などの関係

ない単語に含まれているものも検知してしまい、検知精度を大きく下げている。また、接続詞の「ただ」は想定されていないため、こちらが含まれていた口コミも検知されていた。

#### 1.4 感情分析

口コミから NG ワードを検知するにあたり、投稿者が意図的に NG ワードを使用する際にポジティブな内容で使われることは少ないのではないかと予想した。そこで今回の実験では NG ワード検知の手法に感情分析を取り入れることにした。これにより、NG ワードを含んだ口コミの中でも悪意を持たないものを検知対象から外し、検知確率を向上させることを目的とする。

### 2. 実験方法

本実験では、形態素解析を用いることによって、意図しない単語の使われ方をされている口コミを検知することを目的とする。また口コミに含まれる感情との関係についても実験を行なった。

使用データの概要を表 1 に示す。テストデータについては、NG ワードがそのまま含まれている口コミをさらに目視で判断することによって作成した。

表 1 使用したデータの概要

	引用元	個数
口コミデータ	楽天データセット (楽天市場、みんなのレビュー)	1000件
テストデータ	上記の口コミから目視で判定	1000件
NGワード	ヤマゼンコミュニケーションズ (株) からの提供	101語

<sup>†1</sup> 宇都宮大学 工学部 機械システム工学科 4年  
Utsunomiya University.

<sup>†2</sup> ヤマゼンコミュニケーションズ株式会社  
Yamazen Communications Co.,Ltd.

<sup>†3</sup> 宇都宮大学 工学部 附属ものづくり創生工学センター  
Innovation Center for Research and Engineering Education

## 2.1 検知条件

NGワードを検知する条件として、以下の4つの条件を決定した。

### (1) 単純検知

ロコミの文中にNGワードの文字列が含まれていた場合、このロコミを検知する。この検知方法は従来のシステムと同様のものである。

### (2) 形態素ごとで検知

ロコミのレビューとNGワードをそれぞれ形態素に分解する。その後分解済みのロコミ文に、同じく分解済みのNGワードが同じ順番で含まれていたら、このロコミを検知する。これにより、前述の「いただきます」に含まれている「ただ」など、関係ない別の単語にNGワードの文字列が含まれている場合の誤検知に対応できると考察した。

### (3) 形態素の原形ごとで検知

(2)で行った形態素への分解から、それぞれの形態素の原形を導出する。その後(2)と同様に分解済み原形ロコミ文から分解済み原形NGワードを検知する。これにより「ただ」などの、複数の異なる品詞と意味を併せ持つNGワードのうち片方だけを参照したい時に有効であると考察した。

### (4) 形態素と品詞の組み合わせで検知

(2)で行った形態素への分解から、それぞれの形態素の品詞を導く。その後分解済みロコミと分解済みNGワードが一致、かつそれぞれの品詞も一致した時、このロコミを検知する。NGワードのそれぞれの品詞については、ライブラリのmecabを用いて行なった。しかし、「ただ」という単語については、同社から名詞としてのもののみ検知するよう要求があったため、修正している。

## 2.2 感情分析

本実験において、ロコミの投稿者が意図的にNGワードを用いる時に内容がポジティブな文で使用されることは少ないと予想した。よって、文の内容がネガティブかつNGワードが含まれている時のみに検知することによって、精度が向上すると考えた。

今回の実験では3・1での4つの条件に加え、それぞれネガティブな文のみから判断したものの合計8条件で実験を行った。ロコミ文の感情を判断する為に、一緒に入力された星の数(0~5個、6段階)で判断することにした。今回の実験では、星の数が0~3個のものをネガティブ、それ以外をポジティブとする。

## 3. 結果

今回のプログラムでの検知結果と正解データが一致した確率を表1に示す。この結果より、(1)~(3)に関しては、すべて

表2 検知結果

	全てのロコミ	ネガティブのみ
(1)単純検知	0.897	0.952
(2)形態素ごと	0.939	0.956
(3)形態素の原形	0.915	0.952
(4)形態素と品詞	0.961	0.956

のロコミを対象とするよりネガティブなロコミのみ対象とした方が精度は向上した。また(2)の形態素ごとの検知と、さらに原形に変換した(3)の確率では、(3)の方が確率は減少した。

## 4. 考察

今回のNGワードには「悪い」という単語が含まれているが、文中に「悪くない」という部分が含まれている場合がある。テストデータを作成する際、このような文は投稿者の意図に反するとしてNGワードなしと判断している。しかし、(2)形態素ごとで検知の場合には「悪く」「ない」と分解するので検知できないものの、(3)形態素の原形で検知するには「悪い」「ない」と変換するため、「悪い」の部分で検知してしまう。つまり、NGワードが含まれている動詞や形容詞が否定されている場合、正解とずれてしまう場合がある。今回の実験で(2)より(3)の確率が総じて落ちてしまう原因はこれだと予想される。このような否定語の扱い方に向上の余地があると考えられる。

## 5. 結言

本研究で想定した4つの条件の内3つの条件については、ネガティブなロコミに絞って検知することで検知確率が上昇した。また、すべてのロコミの中から形態素と品詞の組み合わせで検知する方法が最も確率の上昇が大きく、従来の検知方法より6.4%の上昇が確認された。これらはロコミだけでなく、チャットやメールなどにも応用が可能だと考えられる。

将来の展望については、感情分析をSVMなどの分類器を用いて行うことにより、ロコミの「星」の数に依存せずにより検知確率の向上を目指すことを考えている。

### 参考文献

- [1]ITC 総研 “2016年度 SNS 利用動向に関する調査”  
<http://ict.co.jp/report/20160816.html>