

SNSにおける反モラル的な投稿の分析

An Analysis of Cyberbullying Posts in SNS

テーマ：インターネット技術とその応用

指導教員：松本 章代

教養学部 情報科学科

1257247 法井 望美

1. 研究背景

今日、パソコンが普及しインターネットが身近に利用できるようになり生活が便利になった一方、インターネットを利用した人権侵犯事件が後を絶たない [1]。インターネットでは、相手の表情が見えないことから文章のみで解釈・判断しなければならない。また、情報の急速な拡散の可能性を意識して責任を持って情報を扱わなければならない。しかしそのような点を意識して利用している人は少ない。この点に着目し、文章を見る際に、この文章は相手を傷つけてはいないか、不快な文章だと思われないか、と考えさせることによって少しでもネットトラブルを減少させることが必要ではないかと考えた。その第一歩として、本研究では反モラル的な文章とはどのようなものを指すのかを分析する。インターネットでは、ブログ・掲示板・SNS・ホームページ等様々な項目で文章を取り扱っているが、自身が利用していて扱いやすい、クソリブといった相手を不快にさせる返信が目立つ、ツイートの内容から炎上して話題になることがある等、反モラル的な文章が収集しやすいという考えから、SNSの「Twitter(ツイッター)」に的を絞り、「反モラル的な投稿」と「一般的な投稿」にどのような特徴が見られるのかを分析することにした。最終的には、本研究で分析し導き出した反モラル的な投稿の特徴を利用し、反モラル的な書き込みを検知するプログラムを作成する事を目標とする。

2. システム概要

システムは Windows 上で動作しており、開発言語は Ruby を使用する。ツイートの収集は TwitterAPI を用いて自動収集する。日本語の係り受け解析には CaboCha [2] を用いる。また、主成分分析・対応分析には R を使用する。

3. ツイートの調査

3.1 調査対象

炎上が目立ち、クソリブ等反モラル的なツイートが集めやすいと考えられるフォロワー数が多いユーザ [3] を参考に、ユーザのツイートとユーザへのリプライを計 45 のアカウントから 5138 ツイート自動収集する。自動収集したツイートは 3 人でそれぞれ反モラル的だと思ったツイート、一般的だと思われるツイートに手動で分類する。その後 3 人中 2 人が反モラル的だと感じたツイートだけを「反モラル的なツイート」とし、その他は「一般的なツイート」として分類する。結果「一般的なツイート (MT)」は 4405 ツイート、「反モラル的なツイート (AMT)」は 476 ツイート、「反モラル的なリプライ (AMTR)」は 257 ツイートとなった。

3.2 ツイートにおける出現傾向

分類したツイートのそれぞれの割合を見ると、MT は 85.73%、AMT は 9.26%、AMTR は 5.00% となり、このことから Twitter のツイートには比較的モラル的な投稿が多いと考えられる。反モラル的なツイートが多くあると感じるのは、一般的なツイートよりも反モラル的なツイートのほうが人の視線を集めやすく、インターネットの特徴である情報の急速な拡大から炎上という目立つかたちになることが、要因となっているのではないかと推測することができる。

3.3 ツイートにおける形態素の特徴

ツイートは書き手の思うままの文章で表されることが多い。そのため、多種多様な形態素が多く見られた。また、あえて漢字に変換しない平仮名だけのツイートも少なくなく、CaboCha がうまく解析できずに一つの形態素としてカウントする事例もあった。以上のことを考慮し、できるだけ明確に結果が表示されるよう、以下の条件で分析にかけることにする。

- 名詞・動詞を主とした形態素を選択し分析にかける。
- 記号を省略する。例) @, , 「, ~, 等
- ユーザ ID・英単語を省略する。
- 人名・地名を省略する。
- 意味が見いだせない形態素を省略する。例) ジャ、ssbtkkfksww, よれでいいよ

上記の省略の仕方は、指定した範囲の形態素表を目視し手動で削除する。

4. ツイートの形態素の頻度に関する分析

「一般的なツイート」と「反モラル的なツイート」の形態素において、どのような特徴や違いがあるのかを確かめるために、分類した各ツイートの集まりからその頻度に対して分析を行う。この分析によって、各ツイートで使われる形態素の特徴を視覚的にとらえる。主成分分析にあたり、MT、AMT、AMTR を変数、形態素の頻度をケースとする。主成分分析の結果、固有値は第 1 主成分が 1.27、第 2 主成分が 0.91、寄与率は第 1 主成分が 54.57%、第 2 主成分が 27.63% となった。累積寄与率は 82.20% であり、累積寄与率は 70~80% でおおよそが説明できていると受け取れることから、この 2 つの成分で元のデータの多くが説明できているといえる。分析の結果を図 1 に示す。

まず変数に着目すると、MT と AMTR が近い距離にあって、AMT は離れた方向にある。反モラル的なリプライの傾向が一般的なツイートに近いわけとして、相手に対して言葉を発する際は比較的文章に注意して言葉を選んでいないかと推測できる。AMT では頻度が高い順に「きたない」「ばば」「死ね」「ブス」

