

「アトラクションに関する感想情報ツイートの獲得手法」

鈴木絢葉 松井未玖（情報科学ゼミ）

1. はじめに

近年、インターネット上では、ブログ、LINE、Twitter といった様々な SNS が普及してきている。例えば、Twitter のユーザ数は 2015 年 3 月に 3 億 200 万人を超えている[1]。Twitter はユーザ数が増えることにより多くの情報が溢れており、商品やサービスに関する感想などといった評判情報も多くツイートされている。そのため、ツイートから評判情報を抽出するという研究も盛んに行われている[2][3][4]。

商品を購入する前、及び、サービスを利用する前に、何かしらの評判情報を参考にすることがある。例えば、DVD を購入しようとする時、収録されている特典映像の詳細や購入して見た人の感想情報を、Twitter のツイートで獲得し購入を判断することがある。購入の公式ホームページだとおおまかな内容しか記載されていないので、ツイートの感想情報を参考にした方がユーザの選択・決断がよりスムーズになり、安心して買い物することができる。また、通販サイトのレビューには良い感想が書かれる傾向が強いが、Twitter 上の個人の感想には良い感想だけでなく悪い感想も多く含まれているので、商品を購入する上でより有益な感想が得られる。

上記の例のようにツイートの評判情報を利用前に参照することが有益であるものとして遊園地の乗り物が考えられる。遊園地の乗り物は、乗った人でないとその怖さや楽しさが分からないという問題があり、もし乗り物に乗る前に、乗った人の感想を得ることができれば、乗るかどうかの判断に役立つことが期待できるからである。

乗り物に乗った人の感想は、雑誌などの特集記事などにはあるものの、その調査にはコストがかかる。Twitter のツイート情報を利用して、乗った人の感想を収集・分析することができれば、感想情報をコストがかからず収集することが期待できる。そこで本研究では、アトラクションに関する感想情報のツイートを自動抽出する手法を提案する。

2. 提案手法

本手法は、Twitter の検索機能を用いて「乗り物の名前 and (ナガシマ or 長島 or 長スバ or ナガスバ)」というキーワードでツイートを収集されていることを前提とし、検索されたツイートの中から乗り物の感想情報が含まれているツイートを自動的に獲得する手法を提案する。

本手法は、2 段階の処理に分かれている。第 1 段階では、以下の抽出ルールを適用することにより、感想情報が含まれている可能性が高いツイートを獲得する。

■抽出ルール

- ・ 1 文内に「ジェットコースターの名前」と「形容詞」が両方とも含まれている

ただし、文の区切れを示すものとして、句点「。」以外に、「た(助動詞)」及び記号「！」と「？」も文の区切れとして採用した。

第 2 段階では、以下の除外ルールを適用することで、第 1 段階で抽出したツイートから感想情報が含まれていないと思われるツイートを除外する。

■除外ルール

- ・ 抽出ルールを満たす文に「今から」が入っている
- ・ 抽出ルールを満たす文に「ナウ」、「なう」が入っている
- ・ 抽出ルールを満たす文に「～そう」（予想）が入っている

第 2 段階の処理で除外されなかったツイートを感想情報が含むツイートとして獲得する。

3. 実験

提案手法の有効性を確認するために、感想情報ツイートの獲得実験を行った。実験には、対象となるジェットコースターとしてナガシマスパーランドの「アクロバット」を採用した。対象ツイートとしては、Twitter の高度検索により、「アクロバット」を含み、かつ、「ナガシマ」または「ナガスバ」を含むという検索結果で獲得したツイート 200 ツイートとした。この 200 ツイートから感想情報を含むツイートか否かを本手法で判定し、人手で判定した結果を正解データとして、精度と再現率を算出した。実験の結果、精度は 72.9%(35/48)であり、再現率は 44.3%(35/79)であった。

4. まとめ

本研究では、Twitter のツイートからアトラクションに関する感想情報自動抽出し発信するため、Twitter の検索機能を用いて感想情報が含まれているツイートを自動的に獲得する手法を提案した。ツイートの抽出実験の結果、精度は 72.9%(35/48)であり、再現率は 44.3%(35/79)であり、本手法の有効性を確認した。抽出ルールと除外ルールの条件を厳しすぎてしまうと正確な感想情報に絞られると同時にわずかな情報しか獲得できない。いかに正確でより多くの感想情報を獲得し発信できるかが今後の課題である。

参考文献

- [1]「ソーシャルメディアの普及がもたらす変化」、平成 27 年版情報通信白書第 2 部、第 4 章第 2 節、p.199.
- [2]徳田勇介「Twitter からの商品レビュー自動抽出手法の提案」、甲南大学卒業論文、2014.
- [3]中山祐輝、藤井敦、「レビューテキストを対象とした評価条件の抽出手法」、言語処理学会第 19 回年次大会発表論文集、pp.248-251、2013.
- [4]小原基季、森田和宏、泓田正雄、青江順一、「Twitter 本文を用いた観光情報抽出及び分析システムの構築」、人工知能学会全国大会論文集 29、pp.1-3、2015.