

コミックのセリフの感情分析

Sentiment Analysis on Text of Comics

盛 舒峰
Shufeng Sheng

渡辺 裕
Hiroshi Watanabe

早稲田大学大学院基幹理工学研究科
Graduate School of Fundamental Science and Engineering, Waseda University

1. はじめに

感情分析は執筆者の考えがポジティブかネガティブと判断する自然言語処理である。文字情報から間接的に作者の感情をテキストから自動で分析し、出力できるようになれば、作者と読者の意思疎通がよりスムーズになりうる。近年、電子コミック市場の急速拡大は多くが注目を浴びている。したがって、電子化されたコミックのセリフの感情を計算機で自動的に分析することができれば、読者がコミックを読む前に、コミックの登場人物の感情を十分に把握でき、読者が作品をよりよく理解することに役に立つと考えられる。本研究は、コミックのセリフの感情を分析することを目的とする。

2. 提案手法

2.1 Word2vec

まず、900 のコミックのセリフを収集した後、登場人物の発言順に従って整理する。それぞれのコミックのセリフにその感情状態を3種類のパターン（積極的・中性的・消極的（Positive/Neutral/Negative））で分類した。ニューラルネットワークを用いて自然言語処理を行うために、自然言語で記述されたデータを数値に変換する必要がある。本実験では word2vec を使用した。Word2Vec でコミックスのセリフ各単語を、100 次元の対応するベクトル表現に変換した。

2.2 Vader

Vader は、2014 年に提案された感情分析の手法である。ソーシャルメディアに出てくるようなテキストに強い分析力を持っているが、Vader を用いたコミック上の情報の感情分析はまだ実験されていない。本研究では、Vader を利用し、多くのコミックのセリフの感情分析を実行し、感情分析の精度を計算する。

2.3 CNN

畳み込みニューラルネットワーク（CNN）は画像処理系で注目されている手法である。CNN は画像を抽象化する能力を持っているため、実験では CNN によりテキストベクトルでの感情分析を行った。

2.4 RNN

リカレントニューラルネットワーク（RNN）は可変長データをニューラルネットワークで扱う。隠れ層の値を再び隠れ層に入力する構造である。RNN はすでに自然言語処理に幅多く利用されている。実験では RNN を利用してコミックのセリフの感情分析を行った。

2.5 CNN と RNN の併用

コミックの画像にはセリフに含まれない情報があるため、セリフのテキストだけでなく、コミックの画像も考

慮しながら、感情分析を行うことも考えられる。セリフに対応する人物の顔画像から、CNN を利用して顔画像の特徴を抽出し、その特徴ベクトルとテキストベクトル両方を RNN に入力し、感情分析を行う。

2.6 Bi-directional RNN

Bi-directional RNN は、過去の情報だけでなく、未来の情報も考慮することで予測精度を向上させるための手法である。実験ではこの手法も用いてコミックのセリフの感情分析を行った。

3. 結果

今回の実験では、Vader, CNN, RNN, CNN+RNN, BiRNN の五つの手法を利用して、我々が作ったデータセット上で感情分析を行った。その結果を下の表に示す。

表 1. 感情分析の結果

手法	精度(%)
Vader	43.3
CNN	63.9
RNN	70.0
CNN+RNN	67.2
BiRNN	71.1

RNN と BiRNN はそれぞれ 70.0%と 71.1%の精度に達し、Vader, CNN より高い値を示した。また、単純に CNN を使用するのに比べて、CNN と RNN の併用が精度の向上に貢献することが分かった。

4. おわりに

RNN は、自然言語処理上の表現として、Vader や CNN より適すると考えられる。ただし、今回使用したデータセットは小規模であるので、今後はより多くのデータを収集し、豊富な自然言語の特徴を利用して実験する必要があると考えられる。また、コミックのシーン全体的な流れを考慮して、複数のセリフ情報を利用したあるシーンの全体的な感情の分析も検討する。

参考文献

- [1] Tomas Mikolov, Kai Chen, et al.: "Efficient Estimation of Word Representations in Vector Space", In Proceedings of Workshop at ICLR, 2013.
- [2] Hutto, et al: "VADER: A Parsimonious Rule-based Model for Weblogs and Social Media", ICWSM-14,2014.
- [3] Dario Stojanovski, Gjorgji Strezoski, et al.: "Twitter Sentiment Analysis using Deep Convolutional Neural Network", HAIS 2015, 2015.