

審査の結果の要旨

氏名 マレセ テイラ エディソン

この学位請求論文「Mining Fine-grained Opinions in Written and Video Product Reviews with Learned Sentence Representations」では、商品やサービスに対する意見をウェブ上の情報から抽出する意見マイニングにおいて、深層学習の適用を試みるものであり、さまざまな角度から深層学習の手法を適用しその可能性を検証している。

ウェブ上で E コマースや航空券、ホテルなどの購入や予約が一般的になるにつれ、商品やサービスのレビューは意思決定において重要になっている。人々は商品やサービスの購入の意思決定において、こうしたオンラインのレビューを参考にしており、またそれが結果的に良い意思決定につながっているという調査結果も多い。ウェブ上の情報から商品やサービスに対する意見を抽出する技術は、意見マイニングと呼ばれる。意見マイニングのなかでも、より細かい粒度のものに、アスペクトの基づく意見マイニングがある。例えば、ホテルのレビューで、価格は安い、眺望は悪い、サービスは良いなどといったような場合、異なるアスペクトで評価が異なるということになる。したがって、こうしたマイニングにおいては、まずアスペクトを抽出すること、そしてアスペクトごとにユーザが肯定的な意見なのか否定的な意見なのかという感情を分類することの2つが重要になる。いずれのタスクにおいても、品詞タグ付けなどの自然言語処理のアルゴリズムと、Wordnet などの辞書や背景知識を利用する必要があり、どのように人手で素性を作り上げるかが重要な課題になっていた。

一方で、近年の深層学習は、深い階層をもったニューラルネットワークによって表現を学習することができる。特に、オートエンコーダや sequence-to-sequence と呼ばれるモデルは、さまざまな自然言語処理に対しても成果を収めている。そこで、本研究では、深層学習を意見マイニングに適用する。特に、深層学習の特徴を活かした、映像データについての適用も試みる。そのために3つの研究課題を設定している。

まず第一の研究として、従来の研究における再現性の問題を指摘する。従来の研究では、人手で素性を作り込むため、既存研究の精度を再現することが困

難であった。本研究では、既存手法を論文で記述された通りに実装し、論文で報告されている精度が出るかどうかを検証する。その結果、多くの場合、既存手法はハイパーパラメータと呼ばれるアルゴリズム独自のパラメータの変化に鋭敏であり、論文で報告されている精度が出ないことが分かった。同時に、深層学習を用いた手法では、こうしたハイパーパラメータの影響が従来よりも小さく、再現性を確保しやすい可能性があることを示した。

2つ目の研究として、深層学習を用いた意見マイニングを適用する。**Long-short term memory (LSTM)** と呼ばれる手法を用い、アスペクト抽出とアスペクトに基づく感情抽出の両方に適用する。双方向の **LSTM** や、アテンションを用いた手法など、深層学習の分野で知られている手法を組み合わせを用いている。また、再現性を確保するために、さまざまな設定での評価実験を行うことで、意見マイニングに対して、深層学習のアプローチが良好な結果を示すことを実験的に検証した。

3つ目の研究として、映像を対象とした意見マイニングを、深層学習を用いて行う。深層学習は、複数のモダリティ、すなわち、映像と文章などの連携を取りやすいというのが特徴のひとつである。この特徴を活かし、従来は難しかった映像に対する意見マイニングに適用する。音声を抽出し、それとテキストを組み合わせることで、感情抽出に対して高い精度を出せることを示した。

本研究では、意見マイニングという技術分野において、深層学習の適用可能性をさまざまな角度から検証している。特に、従来手法における再現性の問題を指摘したことは重要であり、自然言語処理をはじめとするさまざまな分野で再現性を確保しようという動きはあるが、意見マイニングの分野においても定量的に指摘した。その上で、深層学習の有効性を複数のデータセットで検証している。また、従来にはない深層学習の強みを活かすために、マルチモーダルな情報を用いての意見マイニングが実現できる可能性を示している。以上、本研究は、意見マイニングにおける重要なアプローチを提示しているという点で新規性が高いと同時に、今後、多くの領域で活用できる可能性のある技術であり、その有用性も高い。

よって本論文は博士（学術）の学位請求論文として合格と認められる。