

RGCN と Attention を用いたアスペクトベースの感情分析

樊 惠† 杉本 徹‡

芝浦工業大学大学院 理工学研究科† 芝浦工業大学 工学部‡

1. 研究背景と目的

感情分析は、自然言語処理技術を利用する文書分類タスクの一つである。商品のレビューやツイッターの発言などのコーパスを機械学習あるいはニューラルネットワークのモデルで分析し、一つの文に対して一つの感情極性を予測するのは感情分析の一般的な流れである。しかし、一つの文が一つの感情極性しか持っていないとは限らない。例えば、「The food is delicious, but the service is too bad.」というレビューは二つの感情極性を持つ。従来の感情分析手法は、このようなポジティブとネガティブの両方を含むレビューを分析しにくい。この問題を解決するため、文レベルではなく単語レベルで感情極性を予測するアスペクトベースの感情分析（以下 ABSA）という研究が行われている。ABSA で分析する場合、例文の「delicious」に修飾される「food」の感情極性はポジティブとなり、「bad」に修飾される「service」の感情極性はネガティブとなる。一般的な感情分析に比べると、ABSA の方がより細かく感情極性を分類しており、人間的な分析により近い。しかし、ABSA の研究に使用されるデータは短いレビュー文が多いので、感情の予測には使える情報量が足りないという問題が存在している。そのため、分析対象テキストにおける依存関係などの単語間の関係の情報を取り込むことで情報量を増やすと共に、単語間の関係をモデルに学習させることを試みる。グラフ構造の依存関係を処理するために、本研究はグラフニューラルネットワークモデルを用いる。また、Attention を用い、文中の感情を予測されていない単語（以下コンテキスト）の感情を予測されている単語（以下アスペクトターム）に対する重要度を計算することで、両者の関係もモデルに学習させる。

2. 提案手法

2.1. グラフニューラルネットワーク

本研究は、Graph Convolutional Network[1]（以下 GCN）、Relational Graph Convolution Network[2]（以下 RGCN）と Graph Attention Network[3]（以下 GAT）の三種類のグラフニューラルネットワークモデルを用いる。GCN モデルは各ノードが隣接ノードから特徴量を集約し、自身の特徴量を更新する。これにより、ニューラルネットワークが毎回更新を行う時に、更新された各ノードの特徴量に隣接ノードの特徴量も含まれ、計算に関連する情報が増える。RGCN モデルは GCN モデルの改良型であり、ノードと隣接ノード間の辺の種類を考慮した上で、特徴量を集約することができる。GAT モデルは attention を用いて、各ノードと隣接ノード間の重要度を計算する。この重要度を重みとして特徴量を更新する時に加える。

これらのグラフニューラルネットワークモデルを使うために、ノード特徴量とノード隣接情報が必要となる。

2.2. 単語の埋め込み情報の取得

本研究は、事前学習済みの BERT モデルから単語埋め込みを取得する。データセットによりふさわしい埋め込みを取得するために、事前学習済みの BERT モデルを微調整する。一回微調整した BERT モデルにテキストを入力し、得られるベクトルをノード特徴量として用いる。

2.3. 依存関係の取得

本研究は、Google Natural Language API の依存関係解析器を用いて依存関係の情報を取得する。取得した依存関係の例を図 1 に示す。依存関係にある二つの単語は矢印で繋がっており、矢印の上に依存関係の種類が表示されている。得られる依存関係はノード隣接情報として用いる。

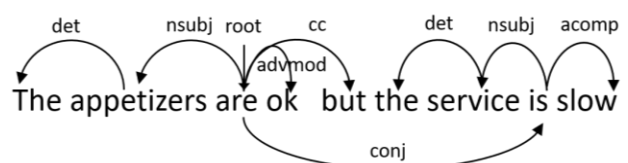


図 1 依存関係の例

2.4. 依存関係の分類

RGCN モデルはノード隣接情報以外に、各辺を区別する情報も必要とする。本研究は、実際にデータセットに出現する 45 種類の依存関係を三つのパターン（修飾、伝送、無関係）に分類し、この三つのパターンで依存関係（辺）を区別する。依存関係の係り先は修飾語である時に「修飾」というパターンに分類する。例えば「delicious←food」間の依存関係「AMOD」は「修飾」に分類する。また、依存関係の係り先は修飾語であるが、係り元は修飾を受ける単語ではない場合もある。例えば「food←is→delicious」の「is」は「delicious」という修飾を受けることができない。「is」は媒介として自分が受け取った情報を「food」に送る。この場合、「food」と「is」間の依存関係は「伝送」に分類する。「修飾」と「伝送」のどちらにも該当しない依存関係は「無関係」に分類する。

2.5. Attention

グラフニューラルネットワークモデルから取得されるテキストの特徴量をコピーし、二つの特徴量を取得する。一つは、アスペクトタームの特徴量に該当する部分を 0 にしてから Attention の input として用いる。もう一つは、コンテキストの特徴量に該当する部分を 0 にしてから Attention の memory として用いる。最後に、input と memory の Attention 係数を計算することで、コンテキストのアスペクトタームに対する重要度をモデルに学習させる。

3. 評価実験

3.1. モデルに関する実験

3.1.1. 実験の方法

モデルに関する実験は、BERT モデルを GCN、RGCN と GAT の三種類のグラフニューラルネットワークモデルと組み合わせ、「BERT」、「BERT+GCN」、「BERT+RGCN」と「BERT+GAT」の四つの手法で行う。ABSA 研究用データセット[4]を用いて四つの手法の適合率、再現率と F 値を計算する。

3.1.2. 結果と考察

モデルに関する実験の結果を表 1 に示す。結果を見ると、「BERT+RGCN」の組み合わせの F 値が最も高い。GCN モデルと比べると、依存関係の種類を区別できる RGCN モデルは、関係の種類に応じてノード特徴量を更新するので、感情分析に関連性が高い依存関係と繋がっているノードの特徴量をより正確に更新していると考えられる。そして GAT モデルの F 値も RGCN モデルに及ばなかった。RGCN モデルは事前に分類された依存関係のパターンを用いて

いるので、依存関係の違いに明確に理解できている。一方 GAT モデルは、自分で依存関係の違いを学習しているので、学習データ量に影響され、依存関係の違いへの理解は RGCN モデルほど深まっていまいと考えられる。

表 1 モデルに関する実験の結果

	適合率	再現率	F 値
BERT	0.6645	0.5768	0.6170
BERT+GCN	0.6940	0.5868	0.6358
BERT+RGCN	0.6817	0.5978	0.6369
BERT+GAT	0.6900	0.5899	0.6360

3.2. Attention に関する実験

3.2.1. 実験の方法

モデルに関する実験で最も良い結果が得られた「BERT+RGCN」モデルに Attention を加えた場合の適合率、再現率と F 値を計算する。そして Attention の有無による結果の変化を考察する。

3.2.2. 結果と考察

Attention に関する実験の結果を表 1 に示す。結果を見ると、Attention を使った方が F 値が高くなる。コンテキストのアスペクトタームに対する重要度の学習は有効であることが示されたと考えられる。

表 2 Attention に関する実験の結果

	適合率	再現率	F 値
BERT+RGCN	0.6817	0.5978	0.6369
BERT+RGCN +Attention	0.6763	0.6262	0.6502

4. まとめと今後の展望

本研究は、RGCN モデルと Attention を用いることで ABSA の精度が向上することを示した。今後は、精度をさらに向上させるために、依存関係の種類を分類する方法をもっと深く研究していく。

参考文献

- [1] Thomas N. Kipf, Max Welling, “Semi-Supervised Classification with Graph Convolutional Networks”, ICLR, 2017
- [2] Michael Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, Max Welling, “Modeling Relational Data with Graph Convolutional Networks”, LNCS, 2017
- [3] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, Yoshua Bengio, “Graph Attention Networks”, ICLR, 2018
- [4] SemEval-2014 Task 4: Aspect Based Sentiment Analysis, Laptop14