

Neural Attention Model を用いた観点付き評判分析

Aspect-based Sentiment Analysis by using Neural Attention Model

柳瀬 利彦 柳井 孝介 佐藤 美沙 三好 利昇 丹羽 芳樹
Toshihiko Yanase Kohsuke Yanai Misa Sato Toshinori Miyoshi Yoshiki Niwa

株式会社 日立製作所 基礎研究センター
Center for Exploratory Research, Hitachi, Ltd.

We propose a neural-network based method for sentiment analysis. In this study, we focus on aspects of sentiment such as prices of food and quality of service in restaurant reviews. Our proposed method consists of two components. The first component is a recurrent neural network that generates a feature vector at each token, and the second one is a neural attention model that recognizes relationship between a given aspect and an expression in a sentence. We empirically confirmed the effectiveness of the proposed method by SemEval 2015 datasets. The accuracies of the proposed method are equivalent to or greater than those of the ranked first team.

1. はじめに

本論文では、観点付き評判分析について議論する。観点付き評判分析とは、例えば、*Pizza here is good, but the service is disappointing.* というレビュー文が与えられた時に、食べ物¹の質 (FOOD#QUALITY) は良いが、サービスの質 (SERVICE#QUALITY) は悪いなど、特定の観点から評判の極性を解析するタスクである。

観点ごとに文を理解すると、従来の評判分析よりも細かい単位 (商品, サービス, 施設等) でユーザの動向を知ることができる。また、この技術は、Argumentation Mining などの他の分野にも応用することが可能である。例えば、議題を観点として議論文を解析すると、議題に対する発言者の賛成・反対の立場を明らかにすることができる [Sato 15]。

本研究では、Neural Network の一種である Neural Attention Model [Luong 15] を用いた観点付き評判分析の方法を提案する。通常の評判分析と異なり本タスクでは、観点对応する文中の表現を特定することが必要である。本手法では、観点と表現の対応を取るために Attention 層を設ける。また、注目する表現の文脈を考慮するために、Recurrent Neural Network (RNN) を用いる。ABSA2015 データセットを用いた評価では、従来の最も高い精度を実現したチームと同等かそれ以上の精度を実現した。

2. 関連研究

言語の意味に関する国際ワークショップ SemEval では、2014 年と 2015 年に Aspect-based Sentiment Analysis サブタスクが開催されている [Pontiki 14, Pontiki 15]。上位の成績を収めているチームの多くは、SVM 等による教師あり学習を採用しており、特徴量には、Bag-of-words (BoW) などの表記に由来するもののほかに、人手で整備された辞書を用いている。そのため、新しい分野に適用する際には、データのほかに辞書の作成も必要になり、コストが大きい。

観点のない一般的な評判分析では、Recurrent Neural Network (RNN) や Convolutional Neural Network (CNN) を用いた手法 [Tai 15, Kim 14] が²、BoW や辞書に基づく手法を

上回る精度を示している。RNN や CNN を用いた手法では、辞書などの人手で作成した言語資源は用いられておらず、新しい分野への適用が容易である。

さらに、観点付き評判分析でも CNN に基づく手法が提案されている [Wang 15]。この手法では、観点と文中の表現の対応を取るため、まず、トークンごとに観点分類を行う。その後、観点を重み付けられた単語ベクトルを、CNN に入力し、極性を判断している。観点分類と、極性解析の 2 段階の推定が必要であるため、観点分類の良し悪しが極性識別の精度に影響を与える。注目箇所³の特定と極性識別を同時に学習する方法が望まれる。

注目箇所³の特定を自動的に行う方法として、機械翻訳では Neural Attention Model が用いられる [Cho 14, Luong 15]。このモデルは Attention 層を持つことが特徴であり、Attention 層は、文脈に合わせて各トークンの重要度を計算する。これにより、文脈にあわせた判断ができる。従来のフレーズベースの翻訳と同等以上の性能を示している。現在では、質問応答 [Hermann 15] や文書要約 [Rush 15] など様々なタスクで使われている。

3. 評判分析のための Neural Attention Model

観点付き評判分析においては、評判を決めるのに先立って、観点对応する文中の表現を認識することが必要である。前述の例では、*Pizza here is good,* が⁴ FOOD#QUALITY に対応し、*but the service is disappointing* が⁵ SERVICE#QUALITY に対応する。

提案手法では、Attention 層が観点と表現の対応をモデル化する。図 1 に、提案する Neural Network の構造を示す。この Neural Network は式 1 の関数 f を表現している。

$$\mathbf{y} = f(x, e, a), \quad (1)$$

ここで、 x は単語系列、 e はエンティティ、 a は属性を表し、 \mathbf{y} は極性 (positive, negative, neutral) の確率である。図 1 では、 x は *Pizza here is good .,*、 e は FOOD、 a は QUALITY である。これらの情報は、単語行列、エンティティ行列、属性行列を参照することで、単語ベクトル \mathbf{x}_t の系列、エンティティベクトル \mathbf{v}_e 、属性ベクトル \mathbf{v}_a に変換される。ここで、単語行

連絡先: 柳瀬 利彦, 株式会社 日立製作所 基礎研究センター,
toshihiko.yanase.gm@hitachi.com

列とは、1行が1つの単語に対応する行ベクトルを持った行列である。エンティティ行列、属性行列も同様である。

\mathbf{x}_t の系列は、双方向RNN (bi-directional Recurrent Neural Network) [Schuster 97]により、前後の文脈を踏まえた、単語の状態ベクトル \mathbf{s}_t に変換される。

まず、通常のRNNについて説明する。RNNは、式2のように、一つ前の状態 \mathbf{h}_{t-1} と入力 \mathbf{x}_t を受け取り、現在の状態 \mathbf{h}_t と出力 \mathbf{s}_t を計算する。

$$\vec{\mathbf{s}}_t = g_s(\mathbf{x}_t, \overleftarrow{\mathbf{h}}_{t-1}), \overleftarrow{\mathbf{h}}_t = g_h(\mathbf{x}_t, \overleftarrow{\mathbf{h}}_{t-1}), \quad (2)$$

ここで、関数 g_s, g_h はRNNのセルを表す。具体的には、Long Short-Term Memory (LSTM) [Sak 14] や、Gated Recurrent Unit (GRU) [Cho 14] がよく用いられる。本研究ではGRUを採用する。

双方向RNNは文を順方向にたどるRNNと、反対方向からたどるRNNからなる。双方向RNNの出力 \mathbf{s}_t は、順方向の出力 $\vec{\mathbf{s}}_t$ と反対方向の出力 $\overleftarrow{\mathbf{s}}_t$ とを結合 (concatenate) したベクトルである。

エンティティのAttention層では、各トークン位置でのエンティティの注目度 ε_t を計算する。まず、トークン位置 t での \mathbf{s}_t と \mathbf{v}_e の関係性 e_t を式3で計算する。

$$e_t = \mathbf{v}_e^T W_e \mathbf{s}_t. \quad (3)$$

次に、式4のように文内での重みの総和が1になるように e_t を変換し、 ε_t を計算する。

$$\varepsilon_t = \frac{\exp(e_t)}{\sum_j \exp(e_j)}. \quad (4)$$

同様に属性のAttention層でも、各トークン位置での属性の注目度 a_t を式5のように計算する。

$$a_t = \mathbf{v}_a^T W_a \mathbf{s}_t, \alpha_t = \frac{\exp(a_t)}{\sum_j \exp(a_j)}. \quad (5)$$

次に、エンティティ・属性のそれぞれの注目度で重み付けされた文ベクトル \mathbf{r} を式6で計算する。

$$\mathbf{r} = \sum_t (\alpha_t \mathbf{s}_t || \varepsilon_t \mathbf{s}_t), \quad (6)$$

ここで、 $||$ は二つのベクトルの結合 (concatenation) を表す。最後に、式7の単層パーセプトロンで \mathbf{y}_p を計算する。

$$\mathbf{y} = \text{softmax}(\tanh(W\mathbf{r} + \mathbf{b})). \quad (7)$$

提案するモデルのパラメータは、バックプロパゲーションにより学習する。これには、RNNやAttention層のほか、単語行列やエンティティ行列、属性行列も含まれる。最適化にはADAM [Kingma 15] と、Stochastic Gradient Descent (SGD) を用いた。学習を安定化するため、勾配の値の最大値を5.0に制限した。また、例外的に長い文による速度低下を避けるため、トークンの最大長を40に制限した。汎化性能を高めるため、Dropout [Srivastava 14] を、RNNの入力と出力に適用し、L2正則化をAttention層と出力層に適用した。Attention層のユニット数は300とした。エンティティ行列は、対象とするエンティティが登場する文の単語ベクトルの平均値を初期値とした。属性行列も、エンティティ行列と同様に初期化した。

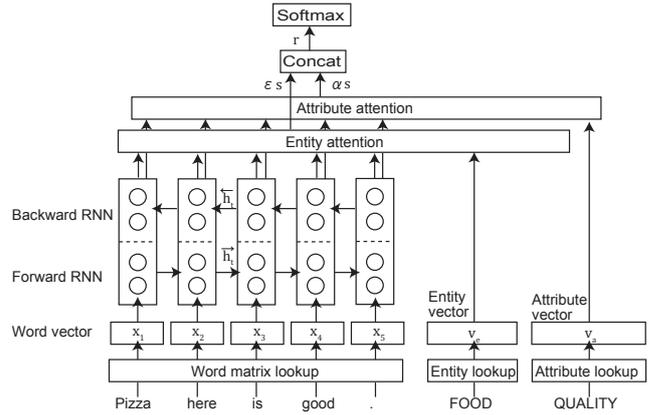


図1: Neural Attention Modelの構造。

4. 評価実験

4.1 実験設定

ABSA15の英語のRestaurantsデータセットとLaptopsデータセットを用いた [Pontiki 15]. Restaurantsデータセットは、訓練データが1315文、テストデータが685文ある。また、Laptopsデータセットは、訓練データが1739文、テストデータが761文ある。メタパラメータ調整のための開発データとして、訓練データのうちの10%を用いた。なお、各文には、複数の観点が割り当てられる場合がある。

ABSA2015のデータセットは、観点なしの評判分析のデータセットであるStanford Sentiment Treebank (SST) [Socher 13]のデータセットと比べるとラベル付きデータの数が少ない。SSTは、文だけでなく、フレーズにも極性が割り当てられており、フレーズ単位では239,232のラベル付きデータが存在する。そのため、Neural Networkのパラメータを十分に学習できない可能性がある。そこで、SSTの全ラベル付きデータを用いて事前学習 (pretraining) を行った。事前学習のラベルは、positive/negativeの二値とした。

前処理として、Stanford Core NLP [Manning 14]により、文からトークン列を得た。Word Embeddingの初期値として、Google News Corpusを生成された300次元のベクトル^{*1}を用いた。Neural Networkの実装には、Tensorflowを用いた [Abadi 15].

実験で用いたメタパラメータの設定を表1に示す。ランダムサンプリングにより、開発データの精度が最も良いものを採用した。なお、RESTはRestaurantsデータセット、LAPTはLaptopsデータセットを表し、(pre)は事前学習ありの場合を表す。Dropout p_k は値を保持する確率である。

4.2 結果

Restaurantsデータセットでの結果を表2に示す。値は、positive, negative, neutralを識別する際の精度 (Accuracy [%]) である。比較として、ABSA2015のベースライン手法と、1位のチームの精度を記載した。ベースライン手法は、Bag-of-wordsを特徴量とする線型SVMによる教師有り学習である [Pontiki 15]. 提案手法は、事前学習を行わない場合でも、ベースラインよりも9ポイント程度優っている。事前学習を行うと、更に7ポイント以上の改善が見られ、ABSA2015の1位のチームの精度を上回った。

*1 <https://code.google.com/archive/p/word2vec/>

表 1: 評価実験に用いたメタパラメタ.

Parameter	REST	REST (pre)	LAPT	LAPT (pre)
Dropout p_k	0.7	0.7	0.7	0.7
learning rate	2.2×10^{-4}	3.8×10^{-4}	1.4×10^{-4}	4.7×10^{-4}
RNN state size	64	256	128	128
minibatch size	16	32	16	16
max epochs	19	12	18	13
L2 coef	8.2×10^{-3}	3.7×10^{-4}	1.0×10^{-3}	9.6×10^{-4}

表 2: ABSA2015 Restaurants データセットでの観点付き極性判定の結果

Method	Accuracy [%]
baseline	63.55
ranked 1st	78.69
proposed	72.07
proposed (pretraining)	79.28

表 3: ABSA2015 Laptops データセット 観点付き極性判定の結果

Method	Accuracy [%]
baseline	69.96
ranked 1st	79.34
proposed	74.60
proposed (pretraining)	79.34

Laptops データセットでの結果を表 3 に示す. Restaurants データセットと同様に, 事前学習をしない場合でも提案手法はベースラインを上回った. また, 事前学習をすることで, 約 5 ポイントの精度改善がみられ, 1 位の手法と同様の精度が見られた.

二つのデータセットで, 提案手法は, BoW や辞書に基づく従来手法と比べて, 同等かそれ以上の性能を示した. さらに, ABSA2015 のデータセットの規模はパラメタを訓練するために十分ではなく, 事前学習によって性能が約 5 から 7 ポイントと大幅に改善することがわかった.

次に, Attention 層による観点と表現の対応のモデル化の有効性について, 事例をもとに調査する. 図 2 は, Attention 層での各トークンのエンティティ, 属性の重みを可視化した際の例である. *Not cheap but very yummy .* という文から, FOOD#PRICES と FOOD#QUALITY の二つの観点について極性を調べている. それぞれの観点について, 上の行がエンティティの重みを表し, 下の行が属性の重みを表す. エンティティは両方とも FOOD であるため, エンティティの重みは共通である. 次に, 属性の重みをみると, FOOD#QUALITY の場合には *yummy* に集中しているのに対し, FOOD#PRICES の場合には *yummy* の重みは小さく, *Not cheap* により大きな重みがかかっている. こうした違いにより, FOOD#PRICES では negative, FOOD#QUALITY は positive と, 異なった判断がされたと考えられる. この例では, 人が見て傾向を解釈できたが, 重みが全体にかかる, 最後のトークンだけに大きな重みがかかるなどの例も多い. これは, RNN によって, 各トークン位置にも周囲の情報が取り込まれているためだと考えられる.

Polarity	Attention	Text
negative	FOOD	Not cheap but very yummy .
	PRICES	Not cheap but very yummy .
Polarity	Attention	Text
positive	FOOD	Not cheap but very yummy .
	QUALITY	Not cheap but very yummy .

図 2: Attention 層を可視化した例.

5. おわりに

本論文では, Neural Network を用いた評判分析について, 特に観点のモデル化に注目して議論した. 本タスクは, 観点のない一般的な評判分析と比べて, 入手可能なデータセットの規模が小さい. データの不足を補うために一般的な評判分析データにより事前学習をすることで, 解決を図った. 評価実験では, ABSA2015 データセットにおいて二つの分野でトップレベルの精度を確認した. 今後の課題として, 日本語を含めた多言語への応用を考えている.

謝辞

本研究を進めるにあたり, 共同研究先である産業技術総合研究所 人工知能研究センターの辻井潤一センター長と中田 亨チーム長には技術的な議論を通して, 貴重なご意見を頂きました. 感謝いたします.

参考文献

- [Abadi 15] Abadi, M., et al.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (2015), Software available from tensorflow.org
- [Cho 14] Cho, K., Merriënboer, van B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y.: Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734, Doha, Qatar (2014), Association for Computational Linguistics
- [Hermann 15] Hermann, K. M., Kočický, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., and

-
- Blunsom, P.: Teaching Machines to Read and Comprehend, in *Advances in Neural Information Processing Systems (NIPS)* (2015)
- [Kim 14] Kim, Y.: Convolutional Neural Networks for Sentence Classification, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746–1751 (2014)
- [Kingma 15] Kingma, D. and Ba, J.: Adam: A Method for Stochastic Optimization, *The International Conference on Learning Representations (ICLR)* (2015)
- [Luong 15] Luong, T., Pham, H., and Manning, C. D.: Effective Approaches to Attention-based Neural Machine Translation, in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1412–1421, Lisbon, Portugal (2015), Association for Computational Linguistics
- [Manning 14] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D.: The Stanford CoreNLP Natural Language Processing Toolkit, in *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55–60 (2014)
- [Pontiki 14] Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., and Manandhar, S.: SemEval-2014 Task 4: Aspect Based Sentiment Analysis, in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 27–35, Dublin, Ireland (2014), Association for Computational Linguistics and Dublin City University
- [Pontiki 15] Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., and Androutsopoulos, I.: SemEval-2015 Task 12: Aspect Based Sentiment Analysis, in *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pp. 486–495, Denver, Colorado (2015), Association for Computational Linguistics
- [Rush 15] Rush, A. M., Chopra, S., and Weston, J.: A Neural Attention Model for Abstractive Sentence Summarization, in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 379–389, Lisbon, Portugal (2015), Association for Computational Linguistics
- [Sak 14] Sak, H., Senior, A. W., and Beaufays, F.: Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition, *INTERSPEECH*, Vol. abs/1402.1128, pp. 338–342 (2014)
- [Sato 15] Sato, M., Yanai, K., Miyoshi, T., Yanase, T., Iwayama, M., Sun, Q., and Niwa, Y.: End-to-end Argument Generation System in Debating, in *Proceedings of ACL-IJCNLP 2015 System Demonstrations*, pp. 109–114, Beijing, China (2015), Association for Computational Linguistics and The Asian Federation of Natural Language Processing
- [Schuster 97] Schuster, M. and Paliwal, K.: Bidirectional Recurrent Neural Networks, *Trans. Sig. Proc.*, Vol. 45, No. 11, pp. 2673–2681 (1997)
- [Socher 13] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C., Ng, A., and Potts, C.: Parsing With Compositional Vector Grammars, in *EMNLP* (2013)
- [Srivastava 14] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research*, Vol. 15, pp. 1929–1958 (2014)
- [Tai 15] Tai, K. S., Socher, R., and Manning, C. D.: Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks, in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 1556–1566, Beijing, China (2015), Association for Computational Linguistics
- [Wang 15] Wang, B. and Liu, M.: Deep Learning for Aspect-Based Sentiment Analysis, *Reports for CS224d, Stanford University* (2015)