

# NAISTにて取り組みたい研究について

## 1 はじめに

### 1.1 NAIST で取り組みたいこと

奈良先端科学技術大学院大学で取り組みたい研究テーマは「SNS から抽出したテキストを対象としたセンチメンタル分析に関する研究」である。本稿では、研究テーマの研究背景・課題、提案内容、研究手法、予想される結果、これまでの修学内容について述べる。

## 2 研究の概要

### 2.1 研究の背景・課題

オンラインで生成されたユーザーコンテンツから意見を発見し、理解することは、多岐にわたる応用にとって重要である。例えば、電子商取引サイトのレビューを分析することで、特定の製品やサービスのどの aspect が高く評価されているか、または改善が必要かを把握できる。しかし、膨大な量のテキストを考慮すると、意見情報を手動で消化するのは非現実的である。そのため、テキストに隠された意見と感情を分析する自動計算フレームワークの設計が必要となり、Aspect-based Sentiment Analysis(ABSA)が注目されている。

普通の感情分析問題は、ターゲットと感情二要素の分析に取り込んでいるが[1]、ABSAにおいては、ターゲットと感情をさらに細かく分けて、メインに4つの要素の抽出を行う。ターゲットは Aspect Category(c) または Aspect Term(a) で表現する。一方、感情は詳細な Opinion Term(t) と Sentiment Polarity(p) で分けられる。これらの要素を単独で抽出する、あるいは複数の要素を一括りにセンチメントタプルとして抽出する、この様に ABSA は様々な種類がある。ASQP(Aspect Sentiment Quad Prediction)で例を挙げると、Input(X) = "Rolls were big, but not good, and sashimi wasn't fresh." という入力の場合、

Output(y) = (Rolls, Food, big, Positive), (Rolls, Food, not good, Negative), (sashimi, Food, n't fresh, Negative) という出力が期待され、それぞれのタプルが要因と感情の関係を示している。

### 2.2 提案内容

最近 Bart、T5 などの LLM は様々な分野で応用され、発展を続けている。text-to-text 問題を解く形で、一回で全ての答えを生成するのが、ABSA の主流となっている[2]。この方法である程度の結果は得たが[3]、seq2seq モデルにおける時系列を考慮した文章生成タスクとは異なり、ABSA のタスクは順序を考慮する必要がなく、未だ改善の余地があると考ええる。前の例で説明すると (Rolls, Food, not good, Negative) が (Rolls, Food, big, Positive) の後にあるのには特別な意味はない。しかし、従来の方法では、学習・推論する際、前のタプルは後ろのタプルの条件となってしまう。さらにタプル単位で見た場合も、a, c, o, s 四つの要素に対しての順序は本来、重要ではない。本研究はこのような生成順序がもたらす影響を消すために、Sota である MvP モデル [4] で使われた Prompt 方法と seq2Path モデル [5] の Path として学習する方法二つを組み合わせ、以上の既存手法モデルと比較して、さらなる性能向上の達成を目指す。

### 2.3 研究手法

モデルの全体的な構成は図1に示してある。

#### 2.3.1 タプル内順序を示す Prompting

モデルにとって、一つの順序は一つの視点と捉えることができる。このように捉えることで、一つの視点という制約を排除し、複数の視点を導入することが考えられる。

このように考えると、異なる順序（四つの要素の全 24 通り）の prompt に基づいて、モデルにさまざまな視点から複数のタプルを生成することができる。また、推論時

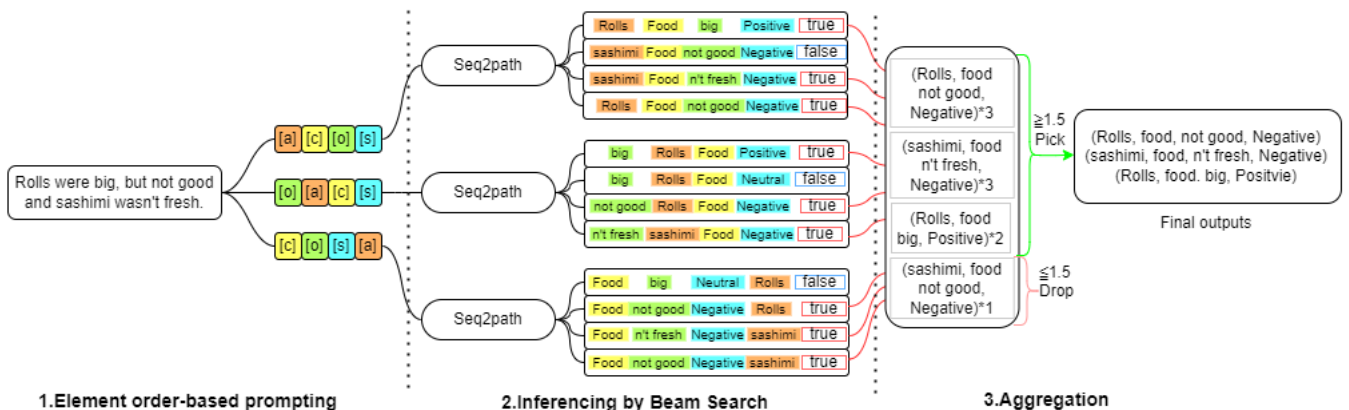


図 1: 本研究で提案するモデルの pipeline

において、一致するタプルを集約し (図1の part3を参考)、半数を超えたタプルを最終結果とすることで、視点単一、あるいは順序が制約となる問題を解決することができる。

具体的な Prompt 方法は MvP を参考に、Input text に”[a][c][o][s]”で示される特殊 token をつける prompting を導入する。例えば、以下になる。

Input (X):Rolls were big. [o][a][c][s]

Output (y): ([o]big, [a]Rolls, [c]Food, [s]Positive)

上のように、Input に対して [o][a][c][s] の順序で prompt を行った場合、その prompt の順番に対応する順序で Output が出力されるように学習される。

### 2.3.2 Pathとしての学習

ABSA において、ある時点での token が推論された際、その次の token を一つに限定するよりも、複数あると考えられる。一つの期待される出力の token から複数の候補を Path として捉えて、学習する手法が Seq2Path から提案された。この手法では、false ケースを取り入れたデータ拡張を行い、その上で、データ構造を Path として捉えることで、Beam Search で推論できるフレームワークを示している。具体的には、図2の左部分を例に挙げると、最初の<bos>token から二つの Path が伸びており、それぞれの<Rolls>と<sashimi>に対して確率が0.5 ずつになる。このように一つの候補ではなく、複数の候補 (Path) を考えて学習を行う。

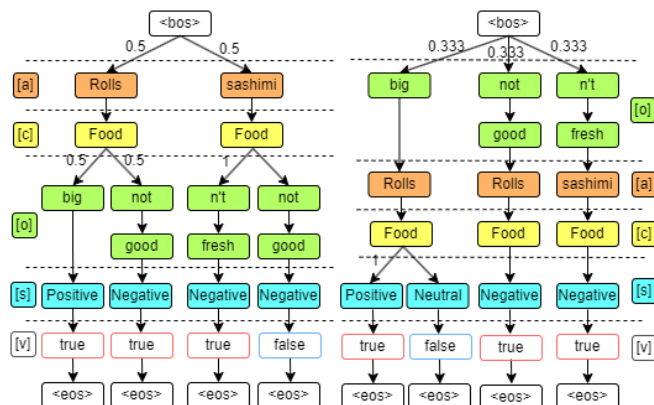


図2: Prompt による Path 構造の違い

### 2.3.3 Prompting+Path 学習

以上の Prompting と Path 学習を組み合わせることを考える。図2のように、Prompting の順番に応じて、出力順序が変わるため、それに合うように Path の接続も変更される。このようにすることで、Prompt を用いた学習を行いながら、Path ライクな学習を可能にする。さらに、この二つを組み合わせることにより、推論時に Beam Search と Aggrigation (図1の Part2 と Part3) を組み合わせることが可能になる。これにより、より信頼された結果を得ることができると考える。実際には、false ケースを示す

([false]) を最終項に導入したデータ拡張を行うことも想定している。

## 3 予想される結果

Seq2Path は 1-to-1 ではなく 1-to-n で学習を行うことができるようになる。しかし、依然として、タプル内の順序に影響を受けてしまう問題があった。MvP は、独自の prompting を導入したことにより、タプル内の順序の影響に対してロバストな学習を行うことができる手法である。自身の提案手法においては、この二つの手法を組み合わせることにより、Path としての学習をさせつつ、タプル内の順序の制約に対処できる仕組みを構築した。タプル内の順番と Path のつながりを連携させることで、二つの手法の良い部分を取り入れることになり、結果として、この二つの手法以上の精度が期待できると考えている。

## 4 これまでの修学内容

学部の卒業論文は四つのセラミック合金材料が送風機アルミニウム製ファンブレードの修復工事への応用に関する研究を行った。アルミニウム材質の特性を考慮し、適切なセラミック合金材料を選び、実験から実用価値の評価まで行うことができた。

卒業した後、Chatgpt から衝撃を受け、AI に興味が湧いて、独学で AI の学習に取り込んだ。自身が四ヶ国語を話せることから、特に NLP に強い関心を持って、基礎的な学習や論文の調査を行っている。

## 参考文献

- [1] B. Liu, “Sentiment analysis and opinion mining,” Synthesis lectures on human language technologies, vol. 5,no. 1, pp. 1–167, 2012.
- [2] Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam.2022. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. CoRR, abs/2203.01054.
- [3] Hang Yan, Junqi Dai, Xipeng Qiu, Zheng Zhang,et al. 2021. A unified generative framework for aspect-based sentiment analysis. arXiv preprint arXiv:2106.04300.
- [4] Zhibin Gou, Qingyan Guo, Yujiu Yang. 2023.MvP: Multi-view Prompting Improves Aspect Sentiment Tuple Prediction. arXiv preprint arXiv:2305.12627
- [5] Yue Mao, Yi Shen, Jingchao Yang, Xiaoying Zhu, and Longjun Cai. 2022. Seq2Path: Generating Sentiment Tuples as Paths of a Tree. ACL 2022, pages 2215–2225, Dublin, Ireland. Association for Computational Linguistics.