

# 係り受け木を考慮するグラフ畳み込みニューラルネットワークによる日本語アスペクトベース感情分析

山口真<sup>1</sup> 狩野芳伸<sup>1</sup>

<sup>1</sup> 静岡大学 情報学部

{syamaguchi, kano}@kanolab.net

## 概要

アスペクトベースの感情分析は、文中の複数のアスペクト語に対する感情極性を特定するタスクである。近年の研究では極性の推定をするために係り受け木を用いてアスペクト語に対するオピニオン語を特定する手法が提案されているが、この手法は英語においてのみ実験されており、日本語における研究では、係り受け木を用いたものはない。そこで本研究では、日本語において係り受け木を用いてアスペクトベースの感情分析を行う手法を提案する。日本語データセットでの感情分析実験により、我々の提案手法が既存の手法よりも優れていることを示す。

## 1 はじめに

アスペクトベースの感情分析は、文中の複数のアスペクト語に対する感情極性を特定するタスクである [1]。例えば「この店のメニューは限られているが、料理は素晴らしい。」という文を考える。この文には2つのアスペクト語「メニュー」と「料理」が含まれており、それぞれの感情極性は「negative」と「positive」となる。一般に、アスペクトベースの感情分析は特定の文とそのアスペクト語に対する感情極性を予測するものとして定式化される。また、感情極性を予測するためには、アスペクト語に対するオピニオン語を正確に特定することが重要である。先の例でいえば、「メニュー」に対しては「限られている」が、「料理」に対しては「素晴らしい」というオピニオン語が対応しており、オピニオン語を特定することでアスペクト語に対する感情極性を推定することが容易になる。近年の研究では、係り受け木を利用してグラフ畳み込みニューラルネットワーク (GCN) [2] を構築することでオピニオン語を効果的に特定する手法が提案され、成果を上げている [3], [4]。しかしこの手法は英語においてのみ実験

されており、日本語での実験は行われていない。そこで本研究では、日本語において係り受け木を用いてアスペクトベースの感情分析を行う手法を提案する。我々の研究の主な貢献は以下の通りである。

- 日本語において係り受け木を用いてアスペクトベースの感情分析を行う手法を提案した。
- 日本語で作成されたデータセットでの実験により、提案手法が既存の手法を上回る性能を示し、本モデルの有効性を示した。

## 2 提案手法

本節では、提案手法の詳細について述べる。提案モデルの全体的な枠組みを図 1 に示す。我々のモデルは先行研究である MGFN [3] を元に日本語に対応するよう拡張したもので、基本的なアーキテクチャは MGFN に準ずる。MGFN モデルは3つの部分から構成されている。1) エンコーディングモジュールは入力分の文脈表現をエンコーディングする。2) グラフ構築モジュールは係り受け解析の結果を用いて潜在グラフを構築する。3) 感情辞書による極性特化モジュールは感情辞書を用いてモデルの注意を感情的により重要な部分に向ける。

本研究では以下の点が MGFN と異なる：

- 係り受け解析器に日本語 UD 解析器である ginza [5] を用いる。
- 日本語で構築された BERT モデルを適用する。また、先行研究と同じ追加学習を行った BERT モデルである Additional pretrained BERT base Japanese finance [6] を適用することで差分を示す。
- 日本語で構築された感情極性辞書である単語感情極性対応表 [7] を用いる。

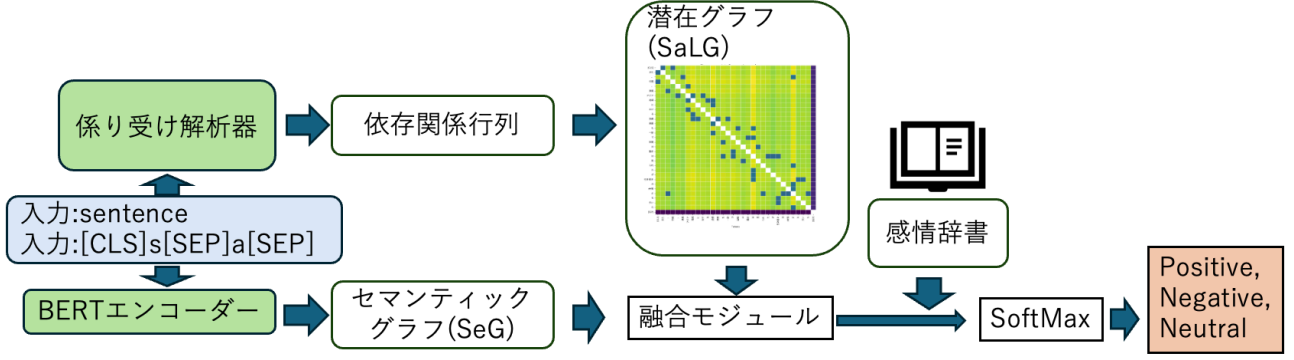


図1 提案手法の全体的な枠組み

## 2.1 エンコーディングモジュール

単語数  $n$  の文  $s = \{w_1, w_2, \dots, w_n\}$  とその中のアスペクト語  $a = \{w_{r+1}, \dots, w_{r+m}\}$  が与えられたとき、文とアスペクト語のペア  $x = ([CLS]s[SEP]a[SEP])$  を事前学習モデルに入力する。出力は  $H = \text{Encoder}(x)$ .  $H = [h_1, h_2, \dots, h_n] \in \mathbb{R}^n \times d$ ,  $h_i$  は単語  $w_i$  の埋め込み表現、 $d$  は隠れ層の次元数である。

## 2.2 グラフ構築モジュール

**セマンティックグラフ** マルチヘッドセルフアテンションによるセマンティックグラフ  $A^{SeG} \in \mathbb{R}^n \times n$  を構築する。

$$A^{SeG} = \frac{\sum_{k=1}^K A^{SeG,k}}{K} \quad (1)$$

$$A^{SeG,k} = \text{softmax}\left(\frac{HW^Q \times (HW^K)^T}{\sqrt{D_H}}\right) \quad (2)$$

ここで  $K$  はアテンションヘッドの数、 $A^{SeG,k}$  は  $k$  番目のアテンションヘッドにおけるセマンティックグラフを表す。  $\sqrt{D_H}$  は文脈表現  $H$  の次元数である。

**係り受けを考慮した潜在グラフ** 入力文に対する係り受け解析の結果を用いて、潜在グラフを構築する。係り受け解析器の結果から係り受け関係行列を作成する。入力文を隣接行列に見立て、係り受け関係のあるノードをエッジで結ぶ。エッジには係り受け関係のラベルを頻度に応じて変換した ID を付与する。この行列をルックアップテーブルとして使い、係り受け関係を単語埋め込み空間に埋め込む。これを用いて構文関係認識行列  $\bar{A} = \text{softmax}(W_a e_{ij} + b_a) \in \mathbb{R}^{n \times n}$  と定義する。  $\bar{A}$  をエッジ重みの初期値として使用し、構文認識潜在木を *tree inducer*[8] を用いて構築する。この潜在木では各ノードが入力文の単語を表す。これによりエッジ重みは対象語句を親とする係り受けによって強化

される。以下を用いて係り受けの根を親とする潜在木のラプラシアン行列  $\hat{L}$  を構築する。

$$\hat{L}_{ij} = \begin{cases} \psi_i + \sum_{i'=1}^n \tilde{A}_{i'j} & i = j \\ -\tilde{A}_{ij} & \text{それ以外} \end{cases} \quad (3)$$

ここで  $\psi_i = \exp(W_r h_i + b_r)$  はノード  $i$  が係り受けの根である確率を表す。  $\hat{L}$  を用いて係り受けを考慮した潜在グラフ  $A_{ij}^{SaLG}$  を以下のように定義する。

$$A_{ij}^{SaLG} = \begin{cases} \tilde{A}_{ij}[\hat{L}^{-1}]_{jj} & i = 1 \text{ and } j \neq 1 \\ \tilde{A}_{ij}[\hat{L}^{-1}]_{ji} & i \neq 1 \text{ and } j = 1 \\ \tilde{A}_{ij}[\hat{L}^{-1}]_{jj} - \tilde{A}_{ij}[\hat{L}^{-1}]_{ji} & i \neq 1 \text{ and } j \neq 1 \\ 0 & i = 1 \text{ and } j = 1 \end{cases} \quad (4)$$

最後に親となる係り受けの根をアスペクト語へするためにルート制約戦略 [9] を用いる。

$$L^r = - \sum_{i=1}^N p_i^r \log \hat{P}_i^r + (1 - p_i^r) \log(1 - \hat{P}_i^r) \quad (5)$$

ここで  $\hat{P}_i^r = \psi_i[\hat{L}^{-1}]_{i1}$  はノード  $i$  が係り受けの根である確率を表し、  $p_i^r \in \{0, 1\}$  はノード  $i$  がアスペクト語である確率を表す。

**グラフの融合**  $A^{SaLG}$  および  $A^{SeG}$  集約関数としてグラフの中間表現を得る。

$$H_l^{SaLG} = \sigma(A^{SaLG} W_c^l H_{l-1}^{SaLG} + b_c^l) \quad (6)$$

$$H_l^{SeG} = \sigma(A^{SeG} W_c^l H_{l-1}^{SeG} + b_c^l) \quad (7)$$

次に2つのグラフを以下により動的に融合する。

$$H_l^{SaLG} = \text{ReLU}(W_l(\alpha H_l^{SaLG} + (1 - \alpha) H_l^{SeG})) \quad (8)$$

$$\alpha = \rho \cdot \sigma(g(H_l^{SaLG})) \quad (9)$$

ここで、  $\alpha$  および  $\beta$  は動的な融合比率を表す。  $g(\cdot)$  は共有畳み込みカーネルを持つ自己ゲーティング関数 [10] である。  $\rho$  は範囲  $[0, 1]$  の先行知識のハイパーパラメータで、  $l \in [1, L]$  はグラフ畳み込みブロックの層数である。

過剰平滑化問題を緩和するために、前層の C-GCN ブロック情報を保持する制御因子  $\omega = \sigma(g(\mathbf{H}_{l-1}))$  を用いる。

$$\mathbf{H}_l^{\text{SaLG}} = \omega \cdot \mathbf{H}_l^{\text{SaLG}} + (1 - \omega) \cdot \mathbf{H}_{l-1}^{\text{SaLG}} \quad (10)$$

SaLG の重要な感情的特徴を抽出するため、潜在特化型注意機構を使用する。

$$\epsilon = \text{softmax}(\mathbf{H}_L^{\text{SaLG}} \mathbf{H}_L^{\text{SeG}}) \quad (11)$$

ここで  $\epsilon$  は、最後の C-GCN ブロックの出力表現に基づくセマンティック認識型潜在重みである。その後、より豊かな感情表現  $z = \epsilon \mathbf{H}_L^{\text{SeG}}$  を得ることができる。また、この特徴をアスペクト語指向にするために、 $z$  に対して、アスペクト語以外にマスクをかけることでアスペクト語指向の感情特徴表現を取得する。

## 2.3 感情辞書による極性特化モジュール

文脈から感情の手がかりを特定するために、感情辞書による極性特化モジュールを導入する。

日本語での実装に際して、入力文の各単語に対して感情スコア  $\eta_k$  を得るために、単語感情極性対応表 [7] を使用した。このスコアを利用して語彙ベクトル  $\text{lex} \in \mathbb{R}^{n \times 1}$  を作成する。ここで、 $\text{lex} = [\eta_1, \eta_2, \dots, \eta_n]$  であり、 $\eta_k$  は  $k$  番目の単語が単語感情極性対応表に存在しない場合は 0 に設定される。同時に、 $l$  層目における  $\mathbf{H}^{\text{SaLG}}$  の隠れ状態表現を、中間ベクトル  $\gamma^{\text{SaLG}} \in \mathbb{R}^{n \times 1}$  にマッピングする。このベクトルは  $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_n]$  で構成され、各低次元ノード表現  $\gamma_k$  は次式によって与えられる：

$$\gamma_k = W^{\text{SaLG}} \mathbf{H}_{l,k}^{\text{SaLG}} + b^{\text{SaLG}} \quad (12)$$

極性特化モジュールの損失関数  $L_s$  を最小化することにより、理想的にはモデルはアスペクト語の感情表現により多くの注意を払うようになる：

$$L_s = (\gamma^{\text{SaLG}} - \text{lex})^2 \quad (13)$$

## 2.4 モデルの学習

ソフトマックス分類器 複数の単語から成るアスペクト語を処理するため、アスペクト語ノードの  $z^A$  に対して平均プーリングを適用し、線形層とソフトマックス関数を使用して感情の確率分布  $\hat{y}^{(s,a)}$  を計算する：

$$\hat{y}^{(s,a)} = \text{softmax}(W^P \text{AvePooling}(z^A) + b^P) \quad (14)$$

ここで、 $(s, a)$  は文とアスペクト語のペアを表す。

学習目標は、以下の全体的な目的関数を最小化する

ることである：

$$L(\Theta) = \lambda L_C + \mu_1 L_r + \mu_2 L_s \quad (15)$$

ここで、 $\Theta$  はモデルのすべての学習可能なパラメータを表す。 $\lambda$ 、 $\mu_1$  および  $\mu_2$  はハイパーパラメータである。主タスクの分類タスクのクロスエントロピー損失  $L_C$  は以下のように定義される：

$$L_C = \sum_{(s,a) \in D} y^{(s,a)} \log \hat{y}^{(s,a)} \quad (16)$$

ここで、 $D$  はすべての文とアスペクト語ペアを含むデータセット、 $y^{(s,a)}$  は感情の真の分布を表す。

## 3 実験設定

### 3.1 データセット

2014 年から 2018 年までの有価証券報告書を用いて作成された「chABSA-dataset」[11] を用いた。データセットにはひとつの文にひとつ以上のアスペクト語が含まれており、それぞれのアスペクト語に対して positive, negative, neutral の 3 つの感情極性のいずれかが 1 名のアノテータにより人手で付与されている。ginza を用いた係り受けと使用した BERT に入力するトークナイザのトークンの不一致があると、グラフ構築モジュールとエンコーディングモジュールとの情報統合ができなくなるため、不一致のある文を除外した。結果、3212 文あったデータセットは 3194 文となった。これを英語版の先行研究同様に訓練データとテストデータに 8:2 で分割した。極性ごとのアスペクト語数の統計を表 1 に示す。

表 1 実験データセットの極性ごとのアスペクト語数

極性	positive	negative	neutral
合計	4289	3091	258
訓練データ	3432	2498	211
テストデータ	857	593	47

### 3.2 提案モデル

入力文に対して係り受けを取得するために UD(Universal Dependencies) 解析器である ja\_ginza\_electra v5.2.0 [5] を用いた。エンコーダーには tohoku-nlp/bert-base-japanese に日本語の金融文書で追加事前学習した BERT モデルである Additional pretrained BERT base Japanese finance [6] を用いた。追加事前学習には 2012 年 10 月 9 日から 2020 年 12 月 31 日までの財務結果の概要及び 2018 年

2月8日から2020年12月31日までの有価証券報告書が用いられており金融コーパスファイルは、約2,700万文で構成されている。バッチサイズは15で訓練され、Adam オプティマイザーを用いて学習率は  $2e-5$  に設定された。GCN の層数は2で設定された。モデルのハイパーパラメータは  $\lambda = 0.5, \mu_1 = 0.05, \mu_2 = 0.05, \rho = 0.2$  と設定された。訓練 epoch の中から損失が最小となったモデルを選択し、テストデータに対して accuracy 及び Macro-F1 を評価した。

### 3.3 ベースラインモデル

先述の Additional pretrained BERT base Japanese finance [6] モデルを訓練データでファインチューニングした。具体的には CLS トークンに加えてアスペクト語の最終層出力同士を加算し、それらを用いてラベルを予測するよう学習した。

## 4 実験結果

### 4.1 総合性能

表2に提案手法とベースラインの性能評価結果を示す。提案手法はベースラインよりも3ポイント程度優れた性能であった。

表2 提案手法とベースライン手法の性能評価結果

モデル	Accuracy	Macro-f1
提案手法	0.9321	0.8783
ベースライン	0.9062	0.8470

### 4.2 アブレーション研究

表3に各損失関数を取り除いた場合のアブレーション研究の結果を示す。提案手法から  $w/o L_r$  を取り除いたほうが性能が良いことがわかる。これはグラフの最上位の親ノードを係り受けの根からアスペクト語にむけるルート制約戦略が、モデルの性能に必ずしも良い影響を与えないことを示している。

表3 アブレーション研究の結果

損失関数	Acc	macro-f1
提案手法	0.9321	0.8783
$w/o L_r$	0.9355	0.8947
$w/o L_s$	0.9335	0.8629

## 5 考察と分析

### 5.1 ケーススタディ

提案手法で正答できるようになった事例の係り受けの解析結果を付録のA.1に示す。この事例について感情辞書による極性特化モジュールの効果を考察するため、感情辞書による極性特化モジュールを取り除く前と後での重みを可視化した(図3)。感情辞書には「欠く」というオピニオン語が negative として登録されているため、この単語が含まれる文は negative として分類されている。GCN を用いたことによりオピニオン語である「欠く」やアスペクト語である「投資需要」以外の予測に必要なないノードの重みが低い、つまり GCN によって適切にプルーニングされたことが可視化結果でみてとれる。

### 5.2 GPT-4o との比較

GPT-4o<sup>1)</sup>に few-shot サンプルを与えた結果との性能比較を表4に示す。与えたプロンプトは付録のA.2に示す。GPT-4o がデータセットの偏りを適切に考慮できなかったため、neutral を破棄したデータセットを用いた。positive, negative のそれぞれのデータ数は同じであり、それぞれについて f1 スコアを計算した。提案手法はこの設定でも GPT-4o を上回る性能を示した。

表4 GPT-4o との比較

モデル	f1-positive	f1-negative
提案手法	0.9452	0.9258
GPT-4o	0.8794	0.8680

## 6 おわりに

本研究では、係り受け木を考慮するグラフ畳み込みニューラルネットワークによる日本語アスペクトベース感情分析を提案した。提案手法は既存手法を上回る性能を示し、感情辞書による極性特化モジュールが効果的であることを示した。本研究も英語版の先行研究同様に長文に対しての適用という課題がある。文中の遠くの単語との関係を考慮するためには GCN において係り受けをより多くホップしながら考慮する必要がある。これがグラフの過平滑化問題とトレードオフの関係にあるため、今後の課題となる。

1) <https://openai.com/index/hello-gpt-4o/>



## 7 謝辞

本研究はJSPS 科研費 (JP22H00804, JP23K22076), JST さきがけ (JPMJPR2461), JST AIP 加速課題 (JPMJCR22U4), およびセコム科学技術財団特定領域研究助成の支援を受けた。

## 参考文献

- [1] Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. SemEval-2014 task 4: Aspect based sentiment analysis. In Preslav Nakov and Torsten Zesch, editors, **Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)**, pp. 27–35, Dublin, Ireland, August 2014. Association for Computational Linguistics.
- [2] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. **arXiv preprint arXiv:1609.02907**, 2016.
- [3] Siyu Tang, Heyan Chai, Ziyi Yao, Ye Ding, Cuiyun Gao, Binxing Fang, and Qing Liao. Affective knowledge enhanced multiple-graph fusion networks for aspect-based sentiment analysis. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, **Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing**, pp. 5352–5362, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics.
- [4] Bingfeng Chen, Qihan Ouyang, Yongqi Luo, Boyan Xu, Ruichu Cai, and Zhifeng Hao. S<sup>2</sup>GSL: Incorporating segment to syntactic enhanced graph structure learning for aspect-based sentiment analysis. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, **Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 13366–13379, Bangkok, Thailand, August 2024. Association for Computational Linguistics.
- [5] 松田寛. GiNZA-Universal Dependencies による実用的日本語解析. 自然言語処理, Vol. 27, No. 3, pp. 695–701, 2020.
- [6] Masahiro Suzuki, Hiroki Sakaji, Masanori Hirano, and Kiyoshi Izumi. Constructing and analyzing domain-specific language model for financial text mining. **Information Processing & Management**, Vol. 60, p. 103194, 2023.
- [7] Hiroya Takamura, Takashi Inui, and Manabu Okumura. Extracting semantic orientations of words using spin model. In Kevin Knight, Hwee Tou Ng, and Kemal Oflazer, editors, **Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL’05)**, pp. 133–140, Ann Arbor, Michigan, June 2005. Association for Computational Linguistics.
- [8] Li Dong, Furu Wei, Chuanqi Tan, Duyu Tang, Ming Zhou, and Ke Xu. Adaptive recursive neural network for target-dependent Twitter sentiment classification. In Kristina Toutanova and Hua Wu, editors, **Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)**, pp. 49–54, Baltimore, Maryland, June 2014. Association for Computational Linguistics.
- [9] Yuxiang Zhou, Lejian Liao, Yang Gao, Zhanming Jie, and Wei Lu. To be closer: Learning to link up aspects with opinions. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, pp. 3899–3909, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [10] Deyu Bo, Xiao Wang, Chuan Shi, and Huawei Shen. Beyond low-frequency information in graph convolutional networks. In **AAAI Conference on Artificial Intelligence**, 2021.
- [11] T. Kubo and H. Nakayama. chABSA: Aspect Based Sentiment Analysis dataset in Japanese. <https://github.com/chakki-works/chABSA-dataset>, 2018. TIS.

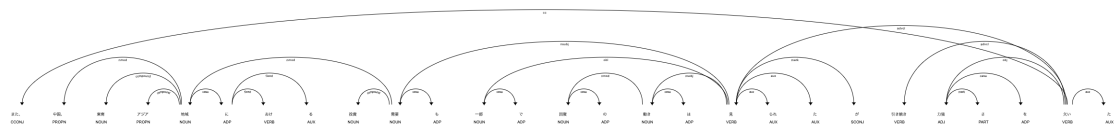


図2 ケーススタディの係り受け解析結果

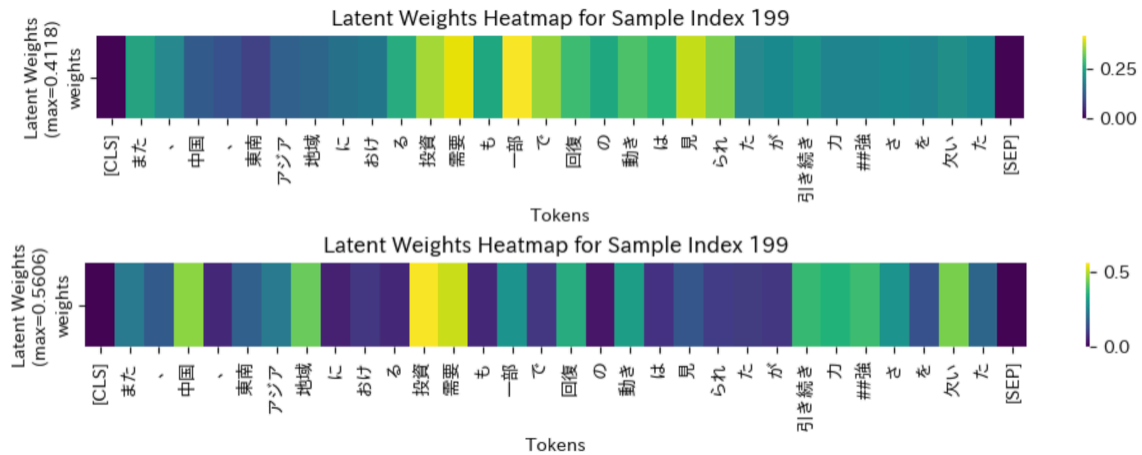


図3 感情辞書を考慮しない場合(上), 考慮する場合(下)

## A 付録

### A.1 ケーススタディの係り受け解析結果

ケーススタディにおける係り受け解析結果を図2に示す。

### A.2 GPT-4o に与えたプロンプト

ここに GPT-4o のプロンプトを記載する。

Sentence: {tokens}  
Target Word: {term}  
Please indicate the sentiment of the targeted word,  
whether positive or not.  
do not output anything other than neutral or not.

また、これに加えて few-shot として以下のようなフォーマットでサンプルを与えた。サンプルは一回のプロンプトにつき各ラベル7つずつ与えた。以下はその一例である。

Sentence: 加えて円高の影響もあり、  
売上高は 19 億 19 百万円（前年比 4.1 %減）  
となりました  
Target Word: その他の事業  
negative