

対話テキスト中の自己主張及び感情の分析に基づく ソーシャルスタイル推定

高橋 憲生 河原 大輔 黒橋 禎夫

京都大学 大学院情報学研究科

{ntakahashi, dk, kuro}@nlp.ist.i.kyoto-u.ac.jp

1 はじめに

これまで、人と人のコミュニケーションを円滑にするために、人の行動に着目した様々な訓練方法や分析方法が提案されてきた。そのような方法として、アサーション (assertion)・トレーニングやソーシャルスタイルがある。アサーション・トレーニングは、自分も相手も大切に自己表現の訓練方法であり、よりよい人間関係の実現に寄与している [1]。ソーシャルスタイルは、人と人のコミュニケーションにおける行動を自己主張表現 (assertiveness) と感情表現 (responsiveness) について分析して得られる 4 つのスタイルである (図 1)[2][3]。対人関係において、各スタイルの特徴に合わせて柔軟に対応することで円滑なコミュニケーションを実現できるため、ビジネスの現場でマネジメントやコーチングに有益な分析方法として利用されている [4]。

テキストを用いた人と人のコミュニケーションにおいても、質問紙法などを用いて、人の行動の結果である発話から自己主張表現と感情表現を分析して、ソーシャルスタイルを推定することは有益である。また、ソーシャルスタイルは、人と機械のコミュニケーションにも適用可能と考えられる。ユーザーのソーシャルスタイルの推定結果をチャットボットなどの自動発話システムに入力することで、ユーザーに合わせて発話を最適化でき、円滑なコミュニケーションを実現できる可能性がある。これまでに感情表現を推定する感情分析の研究は行われてきたが [5]、自己主張表現を推定する研究は行われておらず、ソーシャルスタイルを推定する研究は行われていない。

本研究では、クラウドソーシングでワーカー同士の対話テキストを作成し、対話テキストに自己主張表現と感情表現の強弱をタグ付けする。作成したデータセットから、自己主張表現と感情表現の強弱を推定する二値分類器を作り、ソーシャルスタイルを推定する。

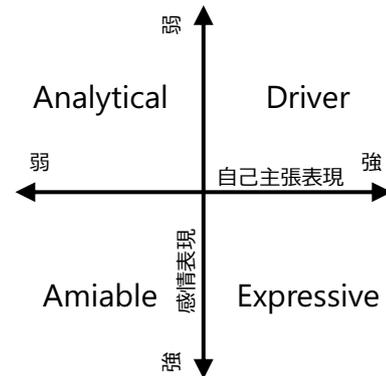


図 1: ソーシャルスタイルの概念図

2 ソーシャルスタイルと自己主張表現

2.1 ソーシャルスタイル

図 1 にソーシャルスタイルの概念図を示す [2][3]。横軸は自己主張表現 (assertiveness) の強弱、縦軸は感情表現 (responsiveness) の強弱を示し、縦軸と横軸で区切られた 4 つの象限は、Driver、Expressive、Analytical、Amiable の各スタイルを示す。ソーシャルスタイルは、他者が客観的に観察可能である行動のパターンであり、内面的なパーソナリティとは異なる。これまでに、IBM Watson Personality Insights に関連する研究 [6] など、テキストからパーソナリティを推定する研究は多数行われてきたが、ソーシャルスタイルを推定する研究は行われていない。本研究では、対話テキストの自己主張と感情の分析から著者のソーシャルスタイルを推定できるか検証する。

2.2 自己主張表現

本論文における自己主張表現 (assertiveness) を定義する。平木は、人間関係における自己表現を、攻撃的自己表現 (aggressive)、非主張的自己表現 (non-assertive)、それらのバランスを取った自己表現 (assertion) の 3 種類に分類しており、自分も相手も大切

| 番号 | 映画タイトル | ワーカー A のコメント | ワーカー B のリプライ | 自己主張 | 感情 |
|-----|------------|-------------------------|-------------------------------|------|----|
| 例 1 | バイオハザード | 飽きずに全作品観ることができるシリーズ作品です | まあね、オイラは渋谷のアンブレラ社に行ってみたいわ !!! | 強 | 強 |
| 例 2 | デスノート | ノートに名前を書かれたらダメ | 犯罪者は確実に減る | 強 | 弱 |
| 例 3 | 幸福の黄色いハンカチ | 果たして黄色いハンカチがあるのかワクワクした | 黄色いハンカチを家の前にかけている場面は印象的だよね | 弱 | 強 |
| 例 4 | 七人の侍 | これぞクロサワ! | ハリウッド映画作品が参考にして いるシーンがわかる | 弱 | 弱 |

表 1: 対話テキストデータセットの自己主張・感情データ作成例 (表中の自己主張と感情はリプライに対するタグ)

にする assertion をよりよい人間関係を築く自己表現として推奨している [1]。ソーシャルスタイルの提唱者である Merrill らは、assertiveness と aggressive には相関があることに言及している [2]。本論文では、自己主張表現 (assertiveness) の定義として、平木の定義を採用する。

テキストを用いたコミュニケーションを理解するうえで、自己主張表現を把握することは有益である。ある論点に対する著者の立場をテキストから推定する研究 [7] や、脅迫や侮辱などの口汚く攻撃的な言葉を検知する研究 [8] など行われてきたが、テキストから自己主張表現の強弱を検知する研究は行われていない。これに対して、Twitter から感情表現を推定する研究 [5] など、感情表現を検知する研究は数多く行われてきた。本研究では、対話テキストデータセットから自己主張表現及び感情表現の強弱を検知する分類器を作成する。

3 対話テキストデータセットの構築

対話テキストデータセットを構築するにあたり、トピックを映画とする。映画を選択した理由は、ほぼ全ての実験協力者が他人の発話を理解可能であり、自らも積極的に発話可能なためである。データセットの作成には、大規模かつ短期間でデータセットを作成可能なクラウドソーシングを利用する。実験協力者 (ワーカー) のソーシャルスタイルを質問紙法で測定した上で、ワーカー同士の対話テキストを作成する。対話テキストから、自己主張表現及び感情表現タグ付きデータ (以下、自己主張・感情データ) とソーシャルスタイルデータから構成される対話テキストデータセットを構築する。表 1 に対話テキストデータセットの自己主張・感情データ作成例を示す。

3.1 ワーカーのソーシャルスタイル測定

質問紙法 [4] を用いて、ワーカーのソーシャルスタイルを測定する。Yahoo!クラウドソーシングを利用してワーカー 100 人を募り、ワーカーは質問紙に回

答する。各ワーカーについて自己主張表現度と感情表現度を求め、ソーシャルスタイルを測定したところ、Driver は 6 人、Expressive は 17 人、Analytical は 39 人、Amiable は 37 人だった。

3.2 対話テキストデータセットの作成

対話テキストの作成方法を以下に示す。

1. ワーカー A に映画タイトルを提示し、ワーカー A がコメントを作成する
2. ワーカー B にワーカー A のコメントを提示し、ワーカー B はリプライを作成する

表 1 は、ワーカー B のリプライについて、自己主張表現の強弱と感情表現の強弱を組み合わせた 4 つの対話例である。平木の定義 [1] に従い、強い自己主張表現は攻撃的自己表現 (aggressive) に近い表現、弱い自己主張表現は非主張的自己表現 (non-assertive) に近い表現とする。強い感情表現は著者の感情をより強く、弱い感情表現は著者の感情をより弱く感じる表現とする。無回答が多い、平均 5 文字程度など極端に発話が短い、相槌のみなど、発話の解析に不適切と判断したワーカー 14 人分の対話テキストは、自己主張・感情データ及びソーシャルスタイルデータから除外する。

次に、ソーシャルスタイルデータの作成方法を示す。各ワーカーのリプライ 10 発話単文について、自己主張表現の強弱と、感情表現の強弱を人手でタグ付けし、ワーカーごとに 10 発話の平均値を計算する。ワーカーごとに、質問紙法の測定結果と、10 発話の平均値からマンハッタン距離¹を計算したところ、極端に距離が離れているワーカーが複数存在した。特に、Analytical と Amiable の自己主張表現度が強くなるケースが多かった。これは、質問紙法は対面コミュニケーションを想定しているが、クラウドソーシングはネット上のテキストを用いたコミュニケーションであり、質問紙法とクラウドソーシングでワーカーが異なる行動を選択している可能性があるためと考えられる²。

¹各座標の差の絶対値の総和

²対面コミュニケーションより、コンピュータを介したコミュニケーションの方が、攻撃性が増す研究報告がある [9]

| スタイル | 人数 | 自己主張表現度 | 感情表現度 |
|------------|-----|---------|-------|
| Driver | 5人 | +0.38 | -0.33 |
| Expressive | 14人 | +0.29 | +0.38 |
| Analytical | 21人 | -0.53 | -0.45 |
| Amiable | 12人 | -0.55 | +0.30 |

表 2: ソーシャルスタイルデータに採用したワーカーの各スタイルの人数、各スタイルの自己主張表現度と感情表現度の平均値

本研究では、質問紙法の測定結果と 10 発話の平均値から推定されるソーシャルスタイルが類似しているワーカーについて、対話テキストから質問紙法の測定結果を推定できるか検証する。マンハッタン距離が 1 以下のワーカーの対話テキストをソーシャルスタイルデータとして採用する。ソーシャルスタイルデータに採用したワーカーについて、質問紙法の測定結果から、各スタイルの人数、各スタイルの自己主張表現度と感情表現度の平均値を集計した結果を表 2 に示す。

最後に、自己主張・感情データの作成方法を示す。今回構築する対話テキストデータセットはワーカーの人数が少なくかつ発話数も少ないため、ソーシャルスタイルを直接推定することは難しい。そこで、マンハッタン距離が 1 より長いワーカーの対話テキストから、自己主張表現と感情表現が特に強いまたは弱い発話を抽出したデータを、自己主張表現と感情表現を二値分類する自己主張・感情データとする。自己主張表現については、攻撃的自己表現 (aggressive) と非主張的自己表現 (non-assertive) のどちらかのタグを手で付けた。感情表現については、感情表現の有りまたは無しタグを手で付けた。

これにより、自己主張・感情データ 794 発話³ (ワーカー 34 人分)、ソーシャルスタイルデータ 3,781 発話 (ワーカー 52 人分) の対話テキストデータセットを構築した。

4 ソーシャルスタイルの推定方法

構築した対話テキストデータセットを用いて、自己主張表現と感情表現の強弱を推定する二値分類器を作り、ソーシャルスタイルを推定する。二値分類するモデルとして、各種タスクで最高レベルの分類性能を実現している BERT [10] を採用する。ワーカーごとに、発話 30~80 文について、一文ずつ自己主張表現と感情表現を二値分類し、分類した結果の平均値を計算し、各ワーカーの自己主張表現度と感情表現度を求め、ソーシャルスタイルを推定する。

³自己主張表現の正例は 394 例で、負例は 400 例。感情表現の正例は 395 例で、負例は 399 例。

5 実験

構築した対話テキストデータセットを用いて、ソーシャルスタイルの推定実験を行う。まず、対話テキストデータセットを用いて、自己主張表現と感情表現について、各々の二値分類器を学習する。次に、ソーシャルスタイルについて、4 章に記載した方法で推定した結果と、質問紙法の測定結果を比較する。

5.1 実験設定

二値分類するモデルとして、日本語 Wikipedia 約 1,800 万文を 30 エポック事前学習した BERT を用いる。BERT のパラメータは英語の BERT_{BASE} と同じである。実験には、自己主張・感情データ 794 発話 (ワーカー 34 人分) とソーシャルスタイルデータ 3,781 発話 (ワーカー 52 人分) を用いる。各発話は、Juman++⁴ を用いて形態素解析し、前処理を行う。二値分類の評価指標には、正答率 (accuracy) を用いる。

5.2 自己主張表現と感情表現の二値分類

自己主張・感情データを用いて、自己主張表現と感情表現それぞれについて二値分類器を学習した。その結果、自己主張表現の正答率 76%、感情表現の正答率 82% の二値分類器を作成できた。

5.3 対話テキストからソーシャルスタイルを推定

表 2 に示す対話テキストデータセットのソーシャルスタイルデータについて、自己主張・感情データを学習したモデルで自己主張表現及び感情表現の二値分類を行い、ワーカーごとに発話の平均値を計算し、ソーシャルスタイル推定結果を得た。図 2 に、モデルの推定結果の散布図を示す。図のマーカーは、質問紙法で分類されたスタイルごと分けた。図 2 に示すとおり、Analytical と Amiable は自己主張表現度が強くなるワーカーが多かったが、推定結果と質問紙法で分類されたスタイルはおおむね一致した。この結果から、対話テキストからソーシャルスタイルを推定することは可能であると考えられる。

表 3 に、質問紙法と推定結果の比較例を示す。例 1 のワーカーは推定結果が正しい例であり、自己主張表現が弱く、感情表現が強い発話が多く見られた。例 2 のワーカーは推定結果が誤っている例であり、文字数が少なく端的で自己主張表現が強い発話が多く見られた。

⁴<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN%2B%2B>

| 番号 | 質問紙法 | | | 推定結果 | | | 典型的な発話例 |
|-----|-------|-------|------------|-------|-------|---------|------------------------|
| | 自己主張 | 感情 | スタイル | 自己主張 | 感情 | スタイル | |
| 例 1 | -0.88 | +0.63 | Amiable | -0.75 | +0.43 | Amiable | ティラノサウルスの迫力がすごく、ドキドキする |
| 例 2 | -0.88 | -0.50 | Analytical | +0.34 | -0.32 | Driver | 頑張れば報われることを証明した |

表 3: 質問紙法と推定結果の比較例

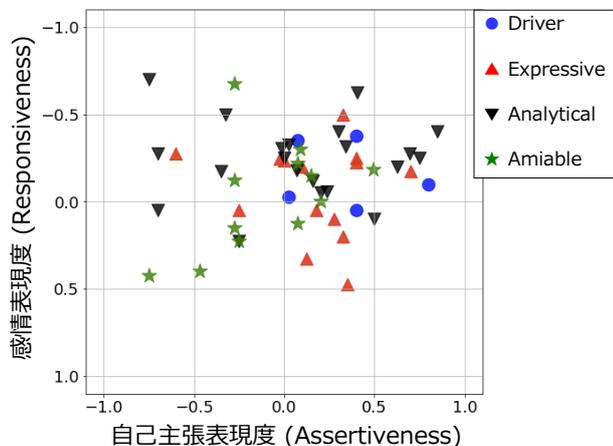


図 2: ソーシャルスタイル推定結果の散布図

5.4 議論

図 2 に示すとおり、推定結果は狭い範囲に集中する傾向があった。これは、各スタイルに特徴的な発話が少なく、スタイルを判別できない発話が多いため、モデル予測の平均値を計算すると各ワーカーに大きな差が生じないためと考えられる。従って、ワーカーの発話を重み付けするなど、推定方法を修正する必要がある。

また、図 2 に示すとおり、推定結果は自己主張表現度が高い傾向があった。自己主張表現を強めにタグ付けしたことが原因の場合は、タグ付け方法を修正する必要がある。他の原因として、3.2 節で言及したとおり、質問紙法の想定状況と、クラウドソーシングのタスク実施状況にずれがあるために発生した可能性がある [9]。その場合は質問紙法をインターネット上のコミュニケーションを想定したものに変更したり、クラウドソーシングの質問内容やタスク設定を変更したりするなど、質問紙法とクラウドソーシングの整合性を図る必要がある。

6 まとめと今後の課題

本研究では、自己主張表現及び感情表現の強弱をタグ付けした対話テキストデータセットを構築した。構築したデータセットから自己主張表現及び感情表現の強弱を検知する二値分類器を作成した。対話テキストの自己主張と感情の分析から著者のソーシャルスタイル

を推定できるか検証し、推定可能である結果を得た。今後の課題は、ソーシャルスタイル推定精度を上げること及び、質問紙法とクラウドソーシングそれぞれが想定するコミュニケーション方法の整合性を図ることである。

謝辞

テキストを用いたコミュニケーションに関する課題を共有していただいた株式会社 cotree の櫻本真理様に感謝いたします。質問紙法 [4] の利用を許諾していただいた室伏順子様にも感謝いたします。

参考文献

- [1] 平木典子. アサーション・トレーニング. 至文堂, 2008.
- [2] David W. Merrill and Roger H. Reid. *Personal Styles and Effective Performance*. CRC Press, 1981.
- [3] Robert Bolton and Dorothy Grover Bolton. *Social Style/Management Style: Developing Productive Work Relationships*. AMACOM, 1984.
- [4] 室伏順子. 苦手なタイプを攻略するソーシャルスタイル仕事術. クロスメディア・パブリッシング (インプレス), 2011.
- [5] Saif Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. SemEval-2018 Task 1: Affect in Tweets. In *SemEval(NAAACL2018 Workshop)*, pp. 1–17, 2018.
- [6] 那須川哲哉, 上條浩一. 日本語における筆者の性格推定の取組み. 言語処理学会第 23 回年次大会発表論文集, pp. 807–810, 2017.
- [7] Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiao-Dan Zhu, and Colin Cherry. SemEval-2016 Task 6: Detecting Stance in Tweets. In *SemEval(NAAACL2016 Workshop)*, pp. 31–41, 2016.
- [8] Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, and Ritesh Kumar. SemEval-2019 Task 6: Identifying and Categorizing Offensive Language in Social Media (OffensEval). In *SemEval(NAAACL2019 Workshop)*, pp. 75–86, 2019.
- [9] 佐藤広英, 日比野桂, 吉田富二雄. CMC(Computer-Mediated Communication) が攻撃性に及ぼす効果. 筑波大学心理学研究, No. 39, pp. 35–43, 2010.
- [10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAAACL2019*, pp. 4171–4186, 2019.