

# 多重対応分析とアスペクトベース感情分析を組み合わせた 受講者満足度調査データの分析手法の開発

藤本一男<sup>1,2</sup> 大畑和也<sup>1,3</sup>

<sup>1</sup> 国立研究開発法人 情報通信研究機構 (NICT) サイバーセキュリティ研究所  
ナショナルサイバートレーニングセンター

<sup>2</sup> 津田塾大学 数学・計算機科学研究所 <sup>3</sup> 法政大学 理工学研究科応用情報工学専攻  
{kazuo.fujimoto,kazuya\_ohata}@nict.go.jp kazuo.fujimoto2007@gmail.com

## 概要

本研究は、NICTが主催するCYDER<sup>1)</sup>の受講者アンケートに対して、多重対応分析(MCA<sup>2)</sup>) [2, 3]とアスペクトベース感情分析(ABSA<sup>3)</sup>) [4]を組みあわせて、概ね「満足」の結果をあらわしているデータから改善すべき課題を抽出する方法の開発である。

本研究では、MCAを用いて、受講生空間を生成しそこに満足度関係の変数を投影することで、満足、および不満の構造を可視化した<sup>4)</sup>。次に、自由記述回答に埋もれている両儀的な語をABSAによって抽出、それをタグ付けし、MCAによって生成された空間上の個体へのマッピングを行なった。こうした処理により全体的に「満足」の回答から課題を抽出することが可能になった。

## 1 はじめに

受講生アンケートにかぎらず、満足度調査データによくみられるのが、全体的に「満足」という傾向である。5件法による選択肢では、殆どの回答が5と4に集中している。しかし、そのイベントは概ね成功であったですませるならともかく、受講者アンケートの目的は、そのイベントの改善にある。であれば、その中からいかに課題を抽出できるかということが重要になる。このような問題設定のもと、我々は、以下の方針で分析を進めた。

- EDAアプローチ [5]によるデータ構造の想定

1) CYDER (Cyber Defense Exercise with Recurrence: 実践的サイバー防御演習)とは、サイバー攻撃を受けた際の一連の対応(インシデント対応)をパソコンを操作しながらロールプレイ形式で体験できる演習 [1]

2) Multiple Correspondence Analysis

3) Aspect Based Sentiment Analysis

4) MCAのこうした使い方は、LeRouxたちによって幾何学的データ解析GDA [2]と呼ばれている。

- 自由記述回答への注目
- アスペクトに基づく自由記述部分の分析
- 選択肢回答分析と自由記述分析の連結

選択肢回答と自由記述分析を組み合わせるアプローチは今日、多く実践されているが、我々のアプローチは、以下の点に独自性を有する。

- 選択肢回答については、MCAを使うことで、個々の回答者の分布位置を幾何学的に確認しながら分析が可能になった。
- 自由回答については、両儀性を有する重要語句に、ABSAを用いpositive/negativeのタグを付与した。
- 各回答者の幾何学的位置と自由回答の関係が分析可能となった。

分析に用いたのは、2020年度に開催されたCYDER, Aコース受講者の受講後アンケート・データ<sup>5)</sup>である。以下、CYDER2020Aとして参照する。

なお、処理言語は、R [6]とPythonを用いた。

## 2 EDAアプローチとMCAによるデータの構造化

EDAアプローチにより変数間の関係を図1のように仮定した。

### 2.1 「スキル向上」と「理解」の関係

まず、変数「スキル向上」が受講がもたらす満足度の指標になること、それを目的変数とすれば、変数「理解」が説明変数となることを確認した。また、その理解-スキル向上関係は、CYDERの実施環境にまつわる変数である「講師説明」「サポート」「対応」「進捗」が関係していることを確認した。

ここで、「理解」と「スキル向上」の関係をみる

5) 受講者総数 2001, 質問総数 (22), そのうち、選択肢回答設問 (15), 自由回答設問 (7)

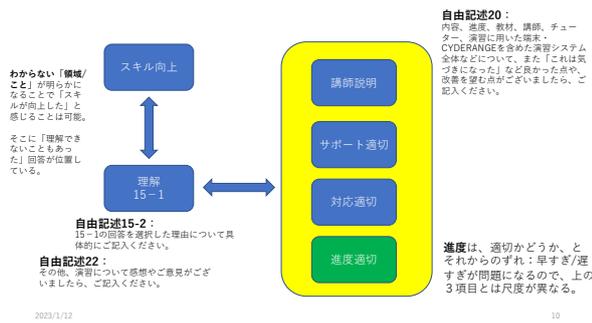


図1 EDAとMCA/CAによるデータの構造化

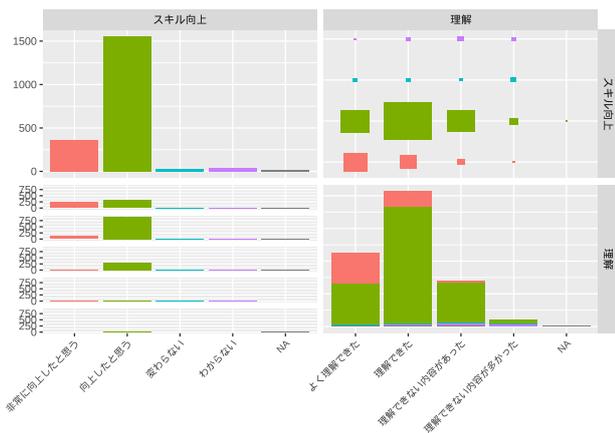


図2 「スキル」と「理解」の対応関係

と、理解の度合いに応じて「スキル向上」感が増加しているわけではないこともあきらかになった(図2)。つまり、理解していることを確認してスキル向上を確認する側面(理解確認によるスキル向上)の他に、自分の理解が欠如していることを自覚することを通じて課題が明確になるという側面での「スキル向上」感もあるわけである(克服課題の明確化というスキル向上)。

## 2.2 自由記述回答への注目

この状況を踏まえて、「理解」の回答を選択した理由を聞く自由回答をテキストマイニングの手法を用いて、分析した<sup>6)</sup>。ここで得られた共起ネットワーク図(図8)をもとにKWIC<sup>7)</sup>で内容を確認すると、頻出語の中に、positiveな用法で用いられているものと、negativeな用法で用いられているものがあることが明らかとなった。この区分は、受講生分析を行う場合に重要であると考えABSAを用いて、自由記述文自体に、Positive/Negativeのタグをつけること、また、この重要頻出語の変数を追加した。

6) 分析ツールとしては、KH Coder [7]を用いた。 <https://khcoder.net/>

7) KeyWord In Context

## 3 ABSAを用いた自由記述文からの意見抽出

サービスや商品提供者はユーザーに対してアンケートやレビューを募ることがあり、ユーザーの意見を理解することはサービスの改善や商品販売の戦略立案につながる。意見を文章で表現することができる自由記述のアンケートやレビューでは、ユーザーは満足もしくは不満足な点についてありのままを記載することができるため、あらかじめ定められた項目や数値を選択する選択肢回答よりもきめ細かな意見の抽出が可能である。自由記述文は数量化して意見を集約する分析がそのままでは難しく、一件ずつ手動で意見抽出を行うことは人件費や時間のコストがかかるため、機械的な分析手法の開発が望まれる。

これらのニーズに対し、深層学習を用いて、自由記述文中に含まれる意見に関する重要な単語を抽出するアスペクトベース感情分析(ABSA)が広く行われている[4]。アンケートやレビューには満足な点と不満足な点が一つの文に混在しているケースがみられ、意見の対象が何であるか、その対象に対してどのような感情を持つかを明らかにするため、アスペクト語の抽出 Aspect Term Extraction (ATE) や抽出した語の感情分析 Aspect Polarity Classification (APC) など、複数の関連タスクが研究されている[8]。

選択肢回答と自由記述の両方を含むアンケートやレビューでは、両方の分析結果を組み合わせることでより詳細な説明性を持つ分析が行えることが期待できる。そこで本研究では、自由記述文に対してABSAを行い、アスペクト語抽出と感情分析から得られた結果を、数量化を行う分析として用いることのできる要素の一つとして加える手法を提案する。

まずはじめに、CYDER2020Aの自由記述文に対してアスペクト語とその感情についてラベル付与をおこない、どの程度アスペクト語の抽出と感情の推定が行えるかを検証する予備実験を行った。その後、残りのデータセットに対しても同様の抽出と推定を行い、図3に示すように、得られた結果を元に出現頻度上位の単語が含まれる回答にフラグを立てることで、選択肢回答の分析で用いることのできる項目として追加した。

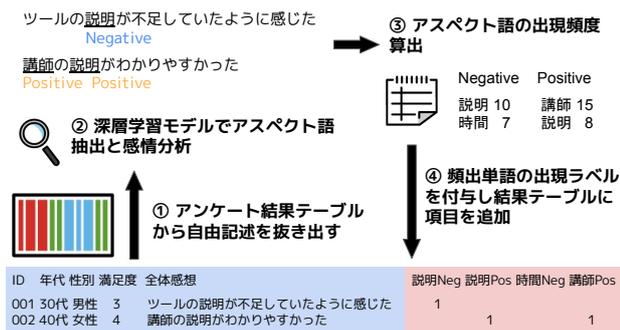


図3 意見抽出と項目追加の枠組み

### 3.1 ABSA モデルの性能評価

ABSA で使用するデータセットは文中に含まれる単語や句ごとの細かい粒度でラベルが付与されるため、一文に対して一つのラベルが付与される一般的な文レベルの感情分析データセットと比較してラベル付与コストがかかる。アスペクト語および感情極性の両方についてラベルが付与された日本語データセットは、我々の知る限り chABSA-dataset<sup>8)</sup> や楽天データセット<sup>9)</sup> と限られている。そのため、本研究ではデータセットのドメインに依存しない一般的な語の抽出を期待して、言語モデル DeBERTaV3<sup>10)</sup> [10] をもとに、英語や中国語を含む8つの言語から20を超えるデータセット<sup>[11]</sup> で学習した多言語モデルを CYDER2020A の自由記述回答に対して適用し、追加の学習を行っていない状態でアスペクト語と感情がどの程度正しく抽出・予測することができるかを予備実験として検証した。

本研究で使用したモデル<sup>9)</sup> は、アスペクト語の抽出と、抽出した単語を positive/Neutral/negative に分類する感情分析を同時に行う、複数のタスクを同時に行うことで性能の向上が期待できるマルチタスク学習<sup>[12]</sup> を導入した LCF-ATEPC<sup>[13]</sup> である。モデルの入力は、文単位で分かち書きした単語シーケンスであり、主な出力はアスペクト語と感情、及び予測確率であるが、本研究ではアンケート結果をもとにした改善と課題抽出を目的としているため、感情が positive または negative と予測されたアスペクト語のみを用い、Neutral と予測された語は全て除外した。

8) <https://github.com/chakki-works/chABSA-dataset>

9) <https://huggingface.co/spaces/yangheng/Multilingual-Aspect-Based-Sentiment-Analysis>

表1 アスペクト語抽出と感情推定のベンチマーク

Model	ATE			APC
	Precision	Recall	F1	Accuracy
LCF-ATEPC	0.449	0.514	0.479	0.981
Random	0.072	0.500	0.126	0.500

### 3.2 予備実験詳細とデータ加工

図1より、CYDER2020A の3つの質問項目に対する自由記述の回答(自由記述 15-2: 理解理由, 自由記述 20: 気づき, 自由記述 22: 全体感想)をそれぞれ文単位に分割し、100件の文に対して手動でラベルを付与した。文の総数は3,350、一文あたりの平均文字数は  $\mu = 31.39$ 、標準偏差  $\sigma = 18.64$  より、文字数  $c$  が  $\mu - \sigma < c < \mu + \sigma$  を満たす文の中から無作為に100件を選び、検証用データとして作成した。データ作成後、各文に対して Mecab<sup>10)</sup> [14] を用いて形態素解析を行い、各トークンごとにアスペクト語であるかどうかと、アスペクト語である場合はその感情ラベル positive/negative を付与したところ、全体で1633個の単語トークンの中に positive 65件, negative 38件のラベルが付与された。なお、図5にラベル付与例とその推論結果を示す。

この検証用データに対して LCF-ATEPC モデル<sup>[13]</sup> を用いて推論したところ、positive 60件, negative 58件, 合計118件のアスペクト語が抽出された。ATE タスクでは各トークンがアスペクト語であるかどうかの一致を Precision, Recall, F1 で評価し、APC タスクでは推論結果と検証用データの両方で抽出された語に対し、positive/negative のタグが一致するかどうかを Accuracy で評価した。表1に、ATE タスクでは予測確率50%で各トークンをアスペクト語として検出するランダムな予測を、同様に APC タスクでは positive/negative 共に50%の確率で予測する、ランダムな予測を1000回行った場合の平均値を Random とし、LCF-ATEPC モデルとの比較を示す。ATE タスクにおいては Precision のスコアが Random と比較して上回り、CYDER2020A データセットが未学習かつ事前学習に含まれない日本語を用いたモデル検証において、ある程度の検出性能を有するといえる。APC タスクにおいては、抽出した単語の感情極性の推論が高い性能で行えることを示しているといえる。

10) <https://taku910.github.io/mecab/>

次に、単語抽出と推論の性能が確認できた LCF-ATEPC モデルを CYDER2020A の残りのデータに対して適用し、アスペクト語の抽出と感情の推論を行い、極性ごとに抽出された単語を出現頻度を基準に並べた。表 2 より、出現頻度上位には、negative/positive 共に「演習」、「内容」、「知識」という単語が共通して含まれていることがわかる。なお、negative で最も出現頻度の多い「時間」は positive ではほとんど抽出されなかった。

表 2 抽出した出現頻度上位 5 単語

Negative		Positive	
アスペクト語	出現頻度	アスペクト語	出現頻度
時間	117	演習	131
演習	85	内容	120
内容	64	説明	90
知識	52	研修	61
用語	40	知識	44

さいごに、自由記述を含むアンケート結果に対して、図 3 に示すように、これら出現頻度上位 5 つの単語及び極性が検出された自由記述文に対してそれぞれ各単語の出現フラグを立てた項目を追加し、選択肢回答分析に用いるデータとして加工した。

#### 4 選択肢回答分析と自由記述分析の連結

こうして、新たに修飾された自由記述部分ももちいて再度、共起ネットワークを確認(図 9)し、「理解」を内容の確認のみならず克服点の確認が自覚されることでもスキル向上を意識できることが明確に区分されていることを確認した。

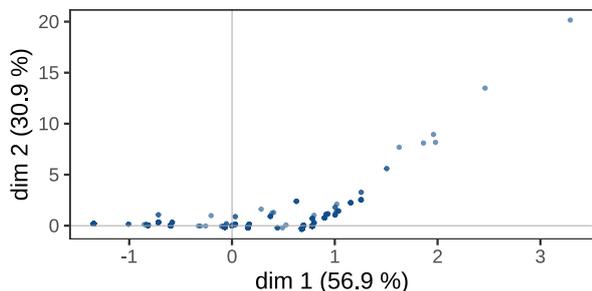


図 4 MCA による個体の分布

以上の過程を経て、MCA によって生成された個体マップ(図 4)<sup>11)</sup>上に、自由回答で得られた重要ポイントを射影することが可能になった(図 7)。この

11) dim1 に沿う形で分布しているポイントが、満足系の回答をした個体である。右側に登っていくラインが確認できるが、こちらが、不満系の回答をした個体の分布である。MCA の場

個体マップは、LeRoux & Rouanet による specificMCA という技法 [2] を用いているが、NA を分析から除去する以外に、低頻度回答(実習に否定的な感想をもった位置の個体)の扱いが問題となった。今回の分析では、全体的に積極的評価がされている中で、そこに埋もれている課題を抽出するのが目的であるので、この低頻度回答も重要だからである。

MCA を用いれば、注目している変数の相互関係を幾何学的に可視化することが可能であり、KH Coder を使えば、自由記述の回答を、回答選択を外部変数として分析することは可能である。

そこで、我々は、両儀的な語に ABSA をもちいてタグ付けを行うことで、自由記述回答の内容を回答者の幾何学的位置を参照しながら分析できる方法を実現した。

回答選択肢の分布は、大多数の回答者が分布する dim1 に沿って密集している「満足」部分と右側にのびていく「不満」部分に区分される。しかし、タグ付けによって明らかになったのは、negative にタグ付された重要語句は、その「満足部分」にも展開されているということであった。この両者の関係を背景として解釈する必要がある。

#### 5 まとめと課題

以上、MCA による幾何学的データ解析(GDA)[2]とアスペクトベース感情分析をもちいた自由記述回答分析を結びつけることで、全体的に満足した反応であるアンケート回答から課題を抽出する方法を開発できた。MCA 用にデータセットを整備していくことで、回答パターンの比較を年次、またコース別に行うことも可能である。

しかし、今回の作業の課程で以下のような課題も明らかになっている。そもそも、自由記述への回答率が低い。その中でも、全体に不満足感想をもつ受講者からの回答がすくない。これを増やすための検討が必要である。また、ABSA について、我々が利用可能な日本語データセットが少ないことから複数の言語で学習した多言語モデルを用いたが、その性能には大いに改善の余地がある。日本語のデータセットで追加学習を行うことで性能の向上が見込まれると考えられるため、ABSA データセット構築が今後の展望として挙げられる。

合は、これらの個体の ID が確認できるので、回答選択肢の内容、自由回答の内容を確認することができる。

## 謝辞

本研究の過程で、CYDER を企画運営している NICT ナショナルサイバートレーニングセンターのみなさんのお世話になりました。記して感謝いたします。

本研究は JSPS 科研費 JP20K02162 「データの幾何学的配置に着目したカテゴリカルデータ分析手法の研究」の助成を受けたものです。

## 参考文献

- [1] NICT. Cyder (cyber defense exercise with recurrence : 実践的サイバー防御演習), 2022. <https://cyder.nict.go.jp/>.
- [2] Brigitte Le Roux, Henry Rouanet, 訳: 大隅昇, 小野裕亮, 鳩真紀子. Multiple Correspondence Analysis(多重対応分析). SAGE publisher (オーム社), 2010(2021).
- [3] Michael J. Greenacre, 訳: 藤本一男. Correspondence analysis in practice Third edition (対応分析の理論と実践). Chapman & Hall/CRC interdisciplinary statistics series. CRC Press, Taylor & Francis Group (オーム社), Boca Raton, 2017(2020).
- [4] Hai Ha Do, PWC Prasad, Angelika Maag, and Abeer Alsadoon. Deep learning for aspect-based sentiment analysis: A comparative review. **Expert Systems with Applications**, Vol. 118, pp. 272–299, 2019.
- [5] John Wilder Tukey. **Exploratory data analysis**. Addison-Wesley series in behavioral science. Addison-Wesley Pub. Co, Reading, Mass, 1977.
- [6] R Core Team. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2022.
- [7] 樋口耕一. 社会調査のための計量テキスト分析-内容分析の継承と発展を目指して、第2版. ナカニシヤ出版, 京都, 2020. OCLC: 1149044718.
- [8] Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. **arXiv preprint arXiv:2203.01054**, 2022.
- [9] 中山祐輝, 村上浩司, Ikuko Hardaway. アスペクトベース意見分析における日本語評価コーパスの構築. 言語処理学会第27回年次大会, 2021.
- [10] Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. Deberta: Decoding-enhanced bert with disentangled attention. In **International Conference on Learning Representations**, 2021.
- [11] Heng Yang and Ke Li. Pyabsa: Open framework for aspect-based sentiment analysis. **arXiv preprint arXiv:2208.01368**, 2022.
- [12] Rich Caruana. Multitask learning. **Machine learning**, Vol. 28, No. 1, pp. 41–75, 1997.
- [13] Heng Yang, Biqing Zeng, Jianhao Yang, Youwei Song, and Ruyang Xu. A multi-task learning model for chinese-oriented aspect polarity classification and aspect term extraction. **Neurocomputing**, Vol. 419, pp. 344–356, 2021.
- [14] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In **Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing**, pp. 230–237. Association for Computational Linguistics, 2004.

## A 参考グラフ

### A.1 自由記述ラベル付与と推論結果比較

ラベル付与結果: 講習はわかりやすかったが画面が小さく見えづかった  
 Positive Negative  
 モデル予測結果: 講習はわかりやすかったが画面が小さく見えづかった  
 Negative  
 ラベル付与結果: 充実した内容で満足したが演習の時間を伸ばしてほしい  
 Positive Negative Negative  
 モデル予測結果: 充実した内容で満足したが演習の時間を伸ばしてほしい  
 Positive Negative  
 ラベル付与結果: ツールの説明が不足していたように感じた  
 Negative Negative  
 モデル予測結果: ツールの説明が不足していたように感じた  
 Negative

図5 ラベル付与の例<sup>†</sup>とモデル推論結果

<sup>†</sup> 個人情報保護の観点より、実際に記載された内容ではなくサンプルとして作成した文を示す。

### A.2 「理解」と「スキル向上」の対応分析

対応分析の出力で、「理解」を標準座標、「スキル」を主座標にした biplot である。矢印が座標軸にあたる。座標軸:「理解できた」「理解できない内容があった」が同じ方向を向き、その方向にスキル「向上したと思う」が位置している。

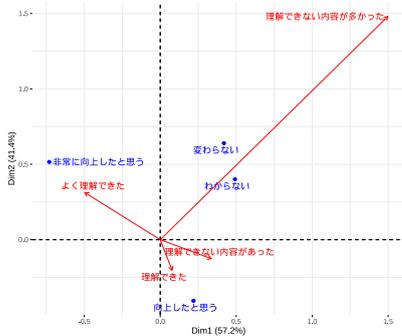


図6 「理解」と「スキル向上」の対応分析

### A.3 ABSA によって「内容」に negative タグをつけた回答者の位置

図6. dim1 は満足度合いの大小 (左が大), dim2 が不満度をあらわしている。自由記述部分の解釈はこれを考慮する必要がある。

### A.4 共起ネットワーク1

図8. この共起ネットワークは、自由記述部分にタグをつけていない。

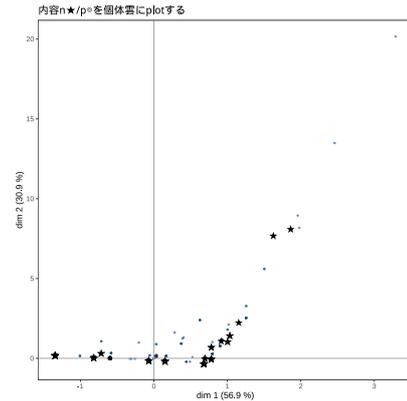


図7 MCAによるタグ付き(内容 negative) 個体の分布

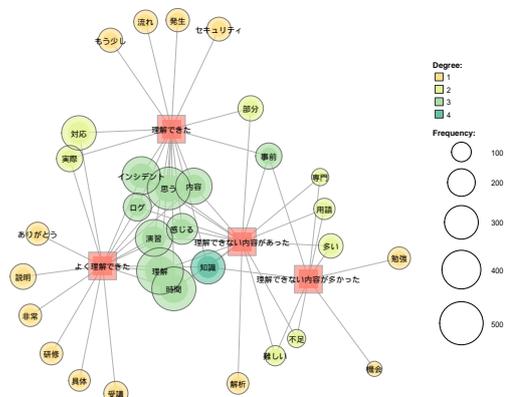


図8 共起ネットワークタグなし

### A.5 共起ネットワーク2

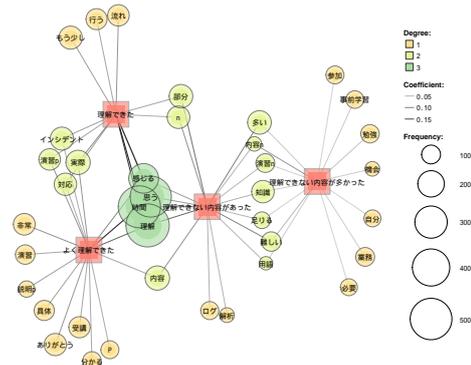


図9 共起ネットワークタグあり

図9. タグをつけ、その語句を強制抽出設定にしたときの共起ネットワーク図