

登場人物と場所の時系列可視化による 物語の出来事の想起支援インタフェース

A Support Interface for Remembering Events in Novels by Visualizing Time-series Information of Characters and their Existing Places

西原陽子^{1*} Jiaxiu Ma² 山西良典³
Yoko Nishihara¹ Jiaxiu Ma² Ryosuke Yamanishi³

¹ 立命館大学情報理工学部

¹ College of Information Science and Engineering, Ritsumeikan University

² 立命館大学大学院情報理工学研究科

² Graduate School of Information Science and Engineering, Ritsumeikan University

³ 関西大学総合情報学部

³ Faculty of Informatics, Kansai University

Abstract: This paper proposes a support interface for remembering events in novels by visualizing time-series information of characters and their existing places. The target users of the proposed interface are people who read multiple novels in parallel. The proposed interface consists of two parts: (a) displaying of novel texts and (b) visualizing of time-series information of character names and their existing places. Given a novel text, the proposed interface extracts character names and places while the Conditional Random Fields (CRFs) is used for the extraction. The proposed interface sets the horizontal axis as the number of the sentence and the vertical axis as characters. The existing places are mapped as labelled dots on the plane given by the two axes. We experimented with participants and verified that the proposed interface could support to remember the events in novels.

1 はじめに

スマートフォンやタブレット、さらに専用の端末 (Amazon kindle など) を用い、電子的に書籍を読む機会が増えた。紙の本と異なり電子書籍は複数冊を所持していても重くないため、携帯が容易にできる。この利点から複数の書籍を並行して読み進めることも可能であり、このような読書形態は今後も増えていくことが予想される。

読書にまとまった時間を確保することが難しく、隙間時間を縫って読書を進める人も多いと考えられる。読書の間隔が空いてしまうと、読書を再開するときに既読部分の内容を忘れてしまい、読み返しをせねばならないことも出てくるだろう。特に多くの人物が登場する物語では、登場人物に関わり起こった出来事が分かりにくくなる。さらに複数の物語を並行して読み進め

ていると、既読部分の内容をより忘れがちになると考えられる。忘れた内容を思い出すための読み返しの量が増えると、読書を先に進めることが難しくなり、結果として読書を楽しむことが難しくなる可能性がある。

そこで本研究では、複数の物語を同時に読み進める人を対象とし、登場人物と場所の時系列情報を可視化することにより、物語の出来事の想起を支援するインタフェースを提案する。物語で記述された順に登場人物と、登場人物がいた場所を可視化することにより、既読部分において登場人物がどこで何をしたか、あるいは何を考えたかを想起しやすくする。これによりスムーズな読書の再開を支援することを目標とする。

長い物語を読む際の読書支援として、登場人物の特徴を示す研究や人物間の関係を示す研究などがある [1, 2]。登場人物の特徴を示す研究として、登場人物の初登場シーンを表示するインタフェースが提案されている [5]。提案インタフェースを用いて読書をするユーザが登場人物名を選択すると、その人物の初出のシーンを表示

*連絡先: 立命館大学情報理工学部
滋賀県草津市野路東 1-1-1
E-mail: nisihara@fc.ritsumeai.ac.jp

する。この研究では登場人物の特徴を失念したときに、すぐに初出のページへと飛び、内容を確認することができるので、素早い読み返しを支援できると考えられる。本研究では本文を読み返すことなく、物語の内容の想起を支援する手法を提案する。

読み返しをさせずに内容の想起を支援する研究として、登場人物間の関係を示すものがある。物語のテキストから登場人物を抽出し、家族関係、仲間関係といった登場人物間の関係を推定するものや [3], ユーザが読書しながらインタフェース上で編集した登場人物の関係を見せるものがある [6]。これらの研究では物語の進捗とともに得られる人間関係やその変化を示すことで、読書の支援を行っている。本研究では物語は登場人物に関わる出来事が順に紡がれてできると考え、既読部分の出来事を示すことにより内容の想起を支援し、読書の支援を行う。出来事はいつ、どこで、誰が、何を、どうしたにより表現されると考え、これらのうちいつ (文番号)、どこで (場所)、誰が (登場人物) の 3 つの情報を抽出し表示することで出来事を示す。

2 提案インタフェース

提案インタフェースの概要を示す。提案インタフェースは物語のテキストを表示する部分と、登場人物と場所の時系列情報を可視化する部分の 2 つにより構成される。提案インタフェースは Web ブラウザ上で動作し、描画には JavaScript のライブラリである D3.js を用いる。

物語のテキストを表示する部分を図 1 に示す。左メニューから読む物語を選択すると、右ウィンドウに物語のテキストが表示される。

登場人物と場所の時系列情報を可視化する部分を図 2 に示す。物語テキストの表示部の右ウィンドウ最上部にある物語のタイトルをクリックすると、時系列情報の可視化結果が表示される。可視化では横軸に文番号を取り、縦軸に登場人物の名前を取る。横軸と縦軸で張られた平面上に人物が居たと考えられる場所がプロットされ、場所名をラベルとして付与する。

可視化図を得るまでのフローチャートを図 3 に示す。入力されるものは物語のテキストである。テキストを文ごとに分割し、文から登場人物と場所表現を抽出する。続いて、文ごとに登場人物と場所表現の紐付けを行い、文番号、登場人物名、登場人物がいた場所の 3 つ組のデータを作成する。3 つ組の情報を基にし、登場人物と場所の時系列情報の可視化を行う。既読部分までの可視化を見ることにより、ユーザは物語内で起こった出来事を人物や場所の情報をヒントとして想起することができると思われる。

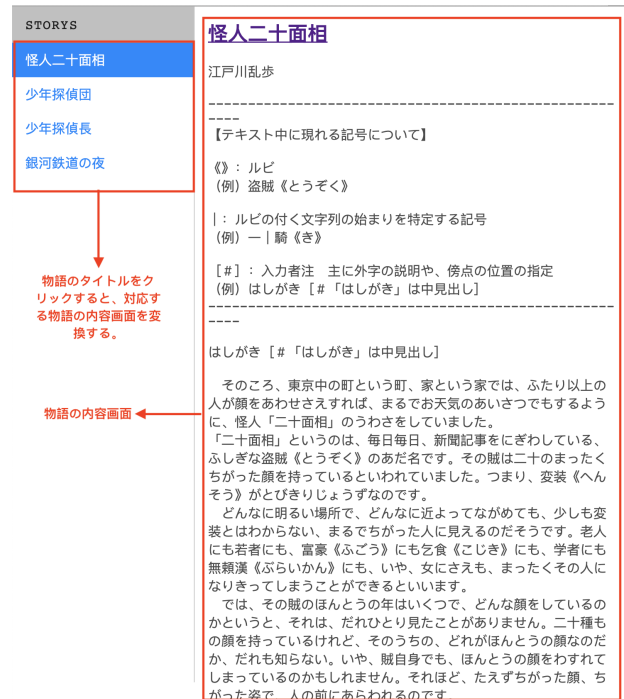


図 1: 提案インタフェースの物語テキストの表示部分

2.1 登場人物と場所表現の抽出

物語のテキストから登場人物と場所表現を抽出する方法を説明する¹。本研究では固有表現抽出に用いられる Conditional Random Fields (CRF) を用い、登場人物と場所表現を抽出する。

はじめに物語テキストを句点 (。や。) により分割し、文の集合を作る。続いて文を形態素解析し、単語と品詞の情報を得る。形態素解析には MeCab、辞書は NEologd を用いる。

続いて人手により各文に含まれる登場人物と場所表現にアノテーションを行う。アノテーションにあたっては BIO2 タグを用いる。BIO2 タグは固有表現抽出に用いられるタグであり、固有表現の先頭の形態素には B タグ、先頭以外の形態素には I タグ、固有表現以外には O タグが付与される。本研究では登場人物と場所表現をそれぞれ分けて抽出を行うため、表 1 に示す 5 種類のタグを用いる。登場人物と場所表現に対するアノテーションを行い、CRF で学習を行うための訓練データを作成する。登場人物は本文に含まれる単語やフレーズのうち、登場人物を明示するものに限定し、代名詞は含まない。同様に場所表現も本文に含まれる単語やフレーズのうち、建物や地名を明示するものに限定し、代名詞は含まない。

訓練データを CRF で学習する際は i 番目の形態素を

¹登場人物と場所表現の抽出については発表済みの内容と同じである [4]。



図 2: 提案インタフェースの登場人物と場所の時系列情報の可視化部分

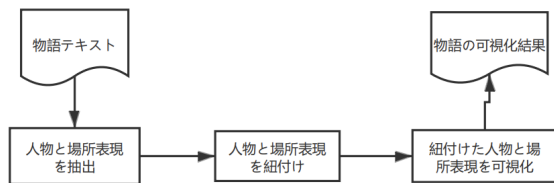


図 3: 可視化図を得るまでのフローチャート

中心とし、前後 3 つの形態素を合わせた合計 7 つの形態素に対する、形態素、品詞情報、BIO2 タグの情報をを用いる。CRF により固有表現抽出器を獲得し、物語テキストからの登場人物と場所表現の抽出を行う。

2.2 登場人物がいる場所の推定

抽出された登場人物と場所表現を紐付け、登場人物がいる場所の推定を行う。一つの文の中に登場人物の表現と場所の表現が含まれる場合、その登場人物と場所には関連があり、登場人物がその場所にいる可能性が高いと考えられる。この仮説の妥当性を示すために予備的な調査を行った。対象とした物語は「桃太郎」「赤ずきんちゃん」「銀河鉄道の夜」であった。

表 1: 人物と場所表現を示す BIO2 タグ

表現タグ	説明
B-CHAR	人物表現文字列の始まり
I-CHAR	人物表現文字列が続いている
B-POS	場所表現文字列の始まり
I-POS	場所表現文字列が続いている
O	人物と場所表現以外の形態素

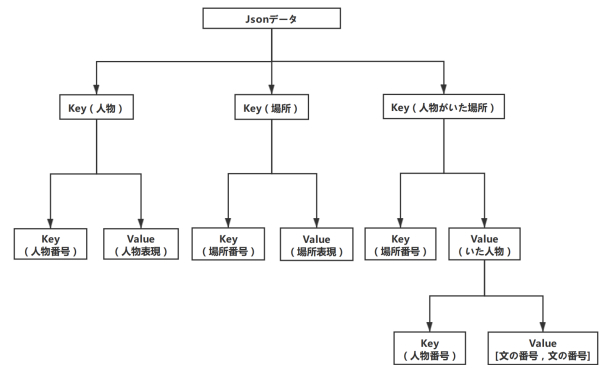


図 4: 提案インタフェースで用いる json データの形式

3 つの物語から登場人物表現のみが抽出されたのは 338 文、場所表現のみが抽出されたのは 73 文、両方が抽出されたのは 99 文、いずれも抽出されなかったのは 322 文であった。両方が抽出された 99 文について登場人物と場所の関連性を手作業にて調べたところ、71% の文で登場人物がその場所にいたことが確認された。このことから一つの文の中に登場人物と場所表現の両方が含まれる場合には、登場人物がその場所に居たと仮定し、紐付けを行うこととする。

登場人物は一文ごとに場所を移動するとは考えにくく、一定の期間、同じ場所に止まると考えられる。そこで本研究ではある文 $s(i)$ においてある登場人物 A とその登場人物 A がいた場所 B が抽出された時、別の文 $s(j)$ (ただし、 $i < j$) において登場人物 A がいた別の場所 C が抽出されるまで、登場人物 A は場所 B に滞在していたとする²。

2.3 登場人物と場所の時系列の可視化

文番号、登場人物、登場人物がいた場所の 3 つ組のデータを受け取り、登場人物と場所の時系列の可視化を行う。3 つ組の情報のデータは json 形式で作成する。図 4 に提案インタフェースで用いる json データの形式を示す。

²この仮説の検証は今後の課題とする。

3 評価実験

電子的に複数の物語を読み進めるユーザに対し、提案インタフェースが既読部分の出来事の想起を支援できるかを評価する実験を行った。

3.1 実験手順

以下の手順で実験を行った。

1. 被験者を3つのグループに分ける。それぞれグループ A, B, C とする。
2. 被験者は1日の間に、指定されたインタフェースで4つの物語の冒頭から指定された部分までを読む。
3. 1週間後に各グループの被験者は物語の出来事に関するテストを受験する。
4. 実験者は得られたテストの平均正解率を各グループで比較し、提案インタフェースの評価を行う。

3.1.1 被験者

被験者は20代の大学生、大学院生で日本人が9名、日本語能力検定試験 N1 の資格を持つ中国人が12名、合計21名であった。全員が日本語の文章を読み、理解する能力を十分に有していた。各グループに日本人を3名、中国人を4名の合計7名ずつ割り当てた。

3.1.2 使用したインタフェース

実験手順の2.で物語を読む際に使用したインタフェースについて説明をする。インタフェースは物語のタイトルと冒頭から指定された部分までのテキストをWebブラウザを通じて表示するものであり、図1で示したものと同一ものを使った。

3.1.3 使用した物語

物語は青空文庫から数回に分けて読むと予想される長編のものを選択した。表2に実験で用いた物語の情報を示す。「少年探偵団」「少年探偵長」「怪人二十面相」「銀河鉄道の夜」の4つを用いた。提案インタフェースのユーザは隙間時間を縫って読書をする想定しており、1回の読書を20分と仮定し、平均的な読書スピードが1分間に400文字であることから積をとり、8,000文字程度で段落の切りの良い部分で区切りをいれた。その結果、読書する文字数は6,000文字から10,000文字までとなった。

表 2: 実験で用いた物語の情報

タイトル	作者	読書指定文字数
少年探偵団	江戸川乱歩	9,979
少年探偵長	海野十三	6,101
怪人二十面相	江戸川乱歩	8,907
銀河鉄道の夜	宮沢賢治	8,903

3.1.4 物語の内容に関するテスト

提案インタフェースでは登場人物と場所の情報を時系列で可視化することにより、物語の出来事の想起を支援するため、人物や場所に関する出来事についての設問からなるテストを作成した。提案インタフェースにより可視化された登場人物と場所の時系列情報を見ることにより、誰がどこで何をしたかを表す出来事の情報想起できたかを確認するテストを作成した。表3に実験で使用したテストの一部を示す。問題は全て物語中の出来事とつながっているものを作成した。例えば問題「ジョバンニは帰る前に何を買った？」は、物語の中で登場人物があるものを買った出来事とリンクしている。人物に関する問題は「1人の人物に関する事」「複数の人物に関する事」であり、場所に関する問題は「1つの場所に関する事」「複数の場所に関する事」とした。複数の人物に関する問題は「Aさんと場所Bにいたのは誰ですか？」が例になり、複数の場所に関する問題は「Aさんは場所Bからどこへ移動しましたか？」などが例になる。1人の人物や1つの場所の問題よりも、複数の人物や場所に関する問題は覚えていなくてはならない人物数や場所数が増えるため、問題としては難度が高くなっている。設問は1つの物語につき5問あり、合計で20問であった。20問のうち12問は記述式、残りの8問は3択の選択肢式であった。テストの受験時間は20問で5分間であった。

テストを受験する際に各グループで差をつけた。グループAはテストを受験する際に、読み返しを一切行わないように指示した。グループBは3分間の読み返しを行った上で、テストを受験するように指示した。グループCは3分間登場人物と場所の時系列が可視化された提案インタフェースを3分間使った上で、テストを受験するよう指示した。

3.2 評価方法

評価は各グループの被験者のテストの平均正解率を用いて行った。グループAとグループBの平均正解率を比較することにより、読み返しが出来事の想起支援に与える影響を明らかにする。続いて、グループAと

表 3: 評価実験用のテスト問題の一部（問題数の括弧内の数字は選択肢式の問題の数）

問題の種類	問題数	問題の例
1人の人物	9(5)	ジョバンニは帰る前に何を買った？(銀河鉄道の夜)
複数の人物	3(1)	春木少年は誰と一緒にカンヌキ山の頂上を登った？(少年探偵長)
1つの場所	5(0)	羽柴氏の宝石はどこに置いている？(怪人二十面相)
複数の場所	3(2)	桂くんはどこから明るい町のほうへかけだした？(少年探偵団)

表 4: 各グループのテストの平均正解率（カッコ内の数字はグループ A からの増減）

グループ	平均正解率
A	26%
B	36%(+10%)
C	54%(+28%)

グループ C の平均正解率を比較することにより、提案インタフェースによる登場人物と場所の時系列の可視化が出来事の想起支援に与える影響を明らかにする。

3.3 実験結果

表 4 に各グループのテストの平均正解率を示す。20 問中の正答数で正解率を出し、各グループの 7 人の平均をとった。提案インタフェースを用いるグループ C の平均正解率が最も高くなった。

問題の種類ごとの平均正解率を表 5 に示す。いずれの問題の種類においても提案インタフェースを用いるグループ C の正解率が最も高くなった。

4 考察

表 4 の全体的な平均正解率を比較すると、グループ A から B にかけての増分は+10%であった。このことから読み返しにより出来事の想起が支援されることがわかった。さらにグループ B からグループ C にかけての増分は+28%であった。この増分は読み返しによる増分よりも多かった。このことから読み返しよりも提案インタフェースを用いる方が、出来事の想起が支援されることがわかった。

続いて人物に関する問題の正解率について考察をする。「1人の人物に関する問題」と「複数の人物に関する問題」の平均正解率を比較すると、「複数の人物に関する問題」に対する方が低かった。この原因としては2つのことが考えられる。1つは、1人の人物についての情報は覚えていても、複数の人物になるとその関連を正確に把握することが難しくなったことがあると考えられる。読み返しや提案インタフェースの使用により関連を把握することができたため、グループ B と C の正解率は増加したと考えられる。もう1つの理由は、選択肢式の問題の割合が挙げられる。「1人の人物に関する問題」は9個中5個が選択肢式であったのに対し、「複数の人物に関する問題」は3個中1個が選択肢式であり、選択肢式の問題の割合がより低かった。このため相対的にグループ A の「複数の人物に関する問題」の平均正解率が低くなり、グループ B や C の増分が多くなったと考えられる。

続いて、場所に関する問題の正解率について考察をする。「1つの場所に関する問題」と「複数の場所に関する問題」の平均正解率を比較すると、「1つの場所に関する問題」に対する方が低かった。この原因としては先と同じく選択肢式の問題の割合が挙げられる。ただ、全てを記述式とした場合にグループ C の正解率は40%となったことから、問題文の選択肢のヒントがない状態で出来事を思い出せるのは高々40%程度となると予想される。今後、クイズの問題を改訂し評価実験を行うことで、人物に関する出来事についても同様の正解率となるかを検証し、提案インタフェースの支援可能性について明らかにしていきたい。

5 おわりに

本論文では複数の書籍を同時に読み進める人を対象とし、登場人物と場所の時系列を可視化することにより物語の出来事の想起を支援するインタフェースを提案した。提案インタフェースは物語のテキストを表示する部分と、登場人物と場所の時系列情報を可視化する部分の2つにより構成される。物語のテキストの各文から登場人物と場所表現を抽出し、文番号と登場人物、登場人物がいた場所の3つ組のデータを作成し、横軸に文番号、縦軸に登場人物名をとった2次元平面に人物がいた場所がプロットされ、場所名を線のラベルとして付与する。提案インタフェースを用いると、既読部分についての出来事の想起が支援され、スムーズな読書の再開の支援が期待される。

提案インタフェースの評価実験では被験者に4つの物語を読ませ、1週間後に物語の内容について問う制限時間付きのテストを実施し、平均正解率を比較した。テストを受験する際に何も使わないグループ A、読み

表 5: 問題の種類ごとの平均正解率 (カッコ内の数字はグループ A からの増減)

グループ	1 人	複数人	1 場所	複数場所
A	40%	14%	6%	33%
B	38% (-2%)	43% (+29%)	20% (+14%)	52% (+19%)
C	57% (+17%)	62% (+48%)	40% (+34%)	62% (+29%)
平均	45%	40%	22%	49%

返しを行うグループ B, 提案インタフェースを用いるグループ C と被験者をグループ分けし, グループごとに平均正解率を比較したところ, 提案インタフェースを用いるグループ C の平均正解率が最も高くなったことから, 既読部分についての出来事の想起を最も支援できることがわかり, 提案インタフェースの有用性を確認した. 今後, テストの問題数と回答形式を揃えた上で評価実験を再度行い, 提案インタフェースの支援の限界について評価をする.

ンコンピュータインタラクション研究会, pp.1-5, (2017).

- [6] 田中翔太郎, 岡部誠, 尾内理紀夫, 物語の内容想起支援インタフェースの開発, 第 55 回プログラミング・シンポジウム予稿集, pp.7-15, (2014).

謝辞

本研究の一部は科研費 (20K12130) によって行われました. 記して謝意を申し上げます.

参考文献

- [1] Apoorv Agarwal, Anup Kotalwar, and Owen Rambow, Automatic Extraction of Social Networks from Literary Text: A Case Study on Alice in Wonderland, Proceedings of the Sixth International Joint Conference on Natural Language Processing, pp. 1202-1208, (2013).
- [2] David Elson, Nicholas Dames, and Kathleen McKeown, Extracting Social Networks from Literary Fiction, Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pp. 138-147, (2010).
- [3] 西原弘真, 白井清昭, 物語テキストを対象とした登場人物の関係抽出, 言語処理学会第 21 回年次大会, pp.628-631, (2015).
- [4] Jiaxiu Ma, 西原陽子, 山西良典, 物語内の人物と場所情報の時系列可視化による読書支援, 第 23 回インタラクティブ情報アクセスと可視化マイニング研究会, pp.9-12, (2019).
- [5] 謝涵, 西田健志, 物語の登場人物を把握しやすくするシステムの提案, 情報処理学会第 174 回ヒューマ

BERT で得られる分散表現に対してコサイン類似度を用いた 小説の挿絵推薦

Recommended Illustrations for Novels Using Cosine Similarity to the BERT based Distributed Representation

下窪 聖人¹ 黄 潤和²

Masato Shimokubo¹, Runhe Huang²

¹法政大学大学院 情報科学研究科

¹Graduate School of Computer and Information Sciences, Hosei University

²法政大学 情報科学部

² Faculty of Computer and Information Sciences, Hosei University

Abstract: In recent years, e-book provides a lot of opportunities that people read books during a commute or break time. E-book reader has some useful functions for reading the books, for instance, Kindle can look up the meaning of words for a reader, and change the font size if necessary. This system enables a function of recommending illustrations during one is reading, which is achieved by calculating cosine similarity to the BERT based distributed representation. There are 4 steps: (1) this system divides text into some scenes by TopicTiling; (2) it extracts the important sentences by BERT for each scene; (3) it extracts named entity using MeCab and CaboCha; and (4) the system recommends appropriate illustrations by calculating similarity between the sentences after this analysis and image title. As expected, this system can understand the contents of the reading easily with recommended illustrations when one is difficult to understand e-book in only text content. The recommended illustrations are original from the illustrations of the web site “irasutoya”. As a result of the experiments, the accuracy of the illustration’s recommendation was about 72.0 percent. This result was obtained by using CNN to convert the image styles of both the original and the recommended illustrations, and then evaluating them by SSIM.

1 はじめに

1.1 研究背景と目的

近年、電子書籍を通勤時間や休憩時間に活用する機会が増えている。このような場面において、読書をしながらか用語を調べたり、知り合い同士で議論し合ったりすることがある。それらをサポートするために、電子書籍リーダーには読書支援ツール機能がある。例えば、小説の単語の意味を調べる機能、テキストを解析して小説の話の流れを可視化する機能などが挙げられる。これらの機能は自然言語処理分野で多く提案されている。しかし、小説を対象とした研究における共通の課題点として、テキストの文脈を高精度で正確に読み取ることが難しい点が挙げられている。この課題点を解決するために多く研究がなされているが、目立った実績はあげられていない。

本研究では、各場面の挿絵を推薦することによって、ユーザの読書をサポートするシステムを提案する。本システムは場面分割をする TopicTiling[1]、重要文を抽出するための BERT[2]を用いる手法[3]、文章解析のための構文解析器 CaboCha[4]と形態素解析器 MeCab[5]を使用し、Web サイト「いらすとや」[6]のイラストを出力する。本システムの出力結果は、シンプルな挿絵を推薦することを目標とする。本研究におけるシンプルな挿絵とは、複数の物体や人物などが描かれている挿絵の場合でも、その挿絵の中で最も重要である物体や人物に注目したイラストと定義する。本システムの有効性を示すことができれば、挿絵の少ない小説に対して挿絵を与えることでユーザの読書をサポートできるようになる。また、書籍の購入時に内容が分からないものがある。そのときに本システムを利用することで、出力から書籍の内容を読み取る手助けになることが期待できる。

1.2 関連研究

読書支援ツールについて様々な研究がある。例えば知的ナビ機能のついた電子書籍リーダーを西岡ら[7]は提案している。この研究では、場面分割の代表的手法である TextTiling[8]と登場人物を対象にした固有表現抽出を使うことで場面を解析する。現在の電子書籍リーダーの読書支援ツールは、単一のキーワードをクエリとした検索は可能であるが、複数のキーワードをクエリとした検索はできないことが課題点として挙げられている。そのため、場面解析を行う提案システムを利用することで課題点を解決できることを示している。またこの課題点を解決することによって、場面ごとに複数の登場人物の出現を検索することが可能になることを示している。

また、村井ら[9]は、ユーザに向けた電子小説の選別を支援するサービス「立ち読み」インタフェースを提案している。このインタフェースは、未読小説のキーワード検索は難しいという課題点を、その小説のレビューとテキスト両方を解析して可視化することで解決している。具体的には、ユーザの興味を引く単語に対して、出現頻度で重要度を求める手法 TF-IDF[10]を用いて興味喚起度を予測し出力するインタフェースである。このとき、単語の興味喚起度を求める手法は、出現頻度が高くなればなるほど、興味喚起度も比例して高くなるようにしている。また、物語の進行における興味を引く度合いの推移を可視化するために、興味喚起度の変化を層として表現している。

推薦システムが普及していく中で Web 小説に關係する推薦システムの例として、小坂ら[11]の研究を挙げる。小坂ら[11]の研究はユーザの作品検索をサポートする推薦システムである。この推薦システムは作品の類似度を示す S スコアと作品の質を示す Q スコアを用いて推薦を行なっている。このときに Doc2Vec[12]で計算する。既存の推薦システムでは解決できていない読者の少ない小説が推薦されにくいという課題点を解決している。Web 小説を用いた実験によって、推薦システムの有効性を示している。

2 提案システム

本システムは、登場人物とその人物の行動を挿絵として示すことでユーザの読書をサポートする。本システムの構成を図1に示す。本システムの処理の流れは、テキストを図1の手順で解析し、その解析結果によって Web サイト「いらすとや」[6]の画像を出力する。小説「浦島太郎」のテキストを入力として与えると図2のようになる。

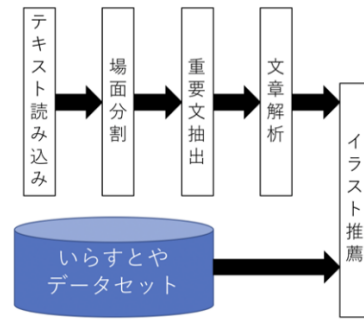


図1 挿絵推薦システムの構成。

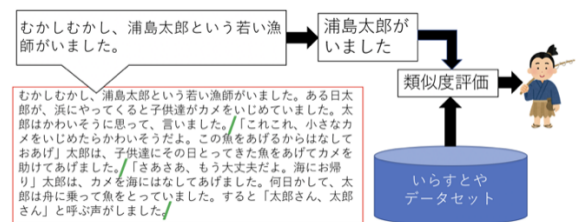


図2 小説「浦島太郎」を入力として与えた結果。

2.1 場面分割

小説や新聞は、読者の理解をサポートするために挿絵を挿入している。挿絵のないテキストのみの小説の挿絵を推薦するためには、テキストを解析して場面を分割し、各場面に対して挿絵を挿入すると読者の理解をサポートできると考える。テキストを場面ごとに分割するときは、作成したベクトルからそれぞれ区切り候補を求める必要がある。一般的に場面を分割するアルゴリズムとして、TextTiling[8]と TopicTiling[1]が使われている。TextTiling[8]は、単語間の類似度と閾値を比較することで分割を行う。この手法の課題点は、1文が短い文章が入力として与えられると、周辺の単語数が少ないことにより適切に場面を分割できないことが知られている。そこで、単語の出現情報を使用しない手法として、トピックモデルを使用する TopicTiling[1]が提案されている。TopicTiling[1]は、トピックモデルの Latent Dirichlet Allocation (LDA) [13]を使用してトピックベクトルを作成し、場面境界の判定に使用する depth score を求めている。LDA[13]は、トピックごとに単語の出現確率の生成や、文章ごとにトピック確率の生成に使用されるモデルである。TopicTiling[1]の depth score は以下の式で求め、 p は位置、 c_p はある位置の coherence score、 $hl(p)$ は p より左側にある範囲で coherence score の最大値を返す関数、 $hr(p)$ は p より右側にある範囲で coherence score の最大値を返す関数である。本システムでは、1文が短い場合にも対

応できる TopicTiling[1]を使用する。また、coherence score はコサイン類似度で求める。

$$\text{depth score} = \frac{1}{2(hl(p) - c_p + hr(p) - c_p)} \quad (1)$$

2.2 重要文抽出

挿絵は各場面の重要な文をイラストとして表していると仮定すると、各場面の文章には内容と関係が薄い文が存在する可能性がある。重要文抽出によってこの文を取り除くことで、読者にとって必要な挿絵を推薦できると考える。重要文抽出は Yang Liu ら[4]が提案した手法を改良して使用する。Yang Liu ら[4]は、初めて文章要約タスクに BERT[3]を適用し、ニュース記事のデータセットで高精度の要約を実現した。Yang Liu ら[4]が提案した手法は英語のみに対応しているため、日本語の小説を処理できるように前処理を加える必要がある。本研究では、形態素解析器 MeCab[5]を使って分かち書きしたテキストを入力として与える。また、日本語の学習済みモデルとして日本語の Wikipedia を学習した京都大学黒橋・村脇研究室のモデル[14]を使用する。

2.3 文章解析

文章解析では、登場人物は主語として必ず出現すると仮定する。このように仮定した理由は、小説は登場人物の行動や発言を中心に話を進める傾向があると考えたからである。本システムの文章解析は、各場面の重要文に対して構文解析と形態素解析を行う。これらの解析により、主語を key、目的語を含む述語を value とする辞書を出力する。例として「うらしまたろうはカメに乗った」を入力すると、辞書 {うらしまたろう : カメに乗った} が出力される。処理の流れとして、まず構文解析で主語と述語を抽出する。処理中にある文の主語が見つからない場合は、直前の主語をその文の主語とした。次に、形態素解析を行い、抽出された主語が適切であるかを判定する。主語が適切でない判定された場合、その key を削除する。そして、形態素解析の処理で残った辞書を出力する。

2.4 イラスト推薦

文章解析後の辞書と Web サイト「いらすとや」[6]のデータセットのラベルとの類似度を比較して、最も類似度が高いラベルを探索する。類似度を求める時には、BERT[2]で得られる分散表現とコサイン類

似度を使用する。「いらすとや」[6]のデータセットから、このラベルを持つイラストを出力する。この出力が本システムの出力結果となる。Web サイト「いらすとや」[6]のデータセットは公開されていないため、スクレイピングによってデータセットを作成した。スクレイピングで収集した情報は、イラストとそのイラストのタイトル（例：石に刺さった剣を抜くアーサーのイラスト）である。Web サイト「いらすとや」[6]のイラストのタイトルには必ず末尾に「のイラスト」と記述されているため、類似度に影響を与える可能性がある。そこで、「のイラスト」の直前までをタイトル（例：石に刺さった剣を抜くアーサー）として保存する。また、Web サイト「いらすとや」[6]には様々なカテゴリが存在している。本研究では、カテゴリ「物語」を対象として 219 枚のイラストを収集した。

3 実験

本実験では、システムの評価を適切に行うため、イラストのある場面のみを対象とし、解析を行った。実験に使用するデータとして、青空文庫[15]で公開されている小説 20 冊を使用する。青空文庫[15]は、著作権が切れた文芸作品などを扱っていることが特徴として挙げられる。実験に使用した青空文庫[15]のテキストの一部を図 3 に示す。イラスト推薦の結果は推薦するイラストと元の挿絵を比較すること、挿絵を推薦する時に使用した文と推薦するイラストのタイトルのコサイン類似度を求めることで評価した。このとき、類似度を算出する前に、CNN[16]を使って画風変換を行うことでスタイルを統一した。元の挿絵との類似度が高ければ高いほど、元の挿絵の情報を多く抽出できているイラストであることを表す。青空文庫[15]のテキストには、[#@@@]を基本形とする[#3字下げ]などのようにレイアウトやルビなどが含まれているため、削除した。また、松本ら[17]が、話し言葉は形態素解析の精度が下がると報告されていたため、会話文の記号である「と」を削除した。

```
[#挿絵] (fig51344_04.png, 横487×縦575) 入る
むかしむかし あるところに 4ひきの こうさが おりました。
なまえは それぞれ
[#5字下げ] フロブシー、
[#4字下げ] モブシー、
[#3字下げ] カトンテル、
[#2字下げ] ピーターです。
4ひきは おかあさんと いっしょに とってもおおきな モミのきの
したにある あなのなかに すんでいました。
```

図3 テキストの一部。

3.1 場面分割の評価方法

TopicTiling[1]での場面一致度は再現率, 適合率, F 値によって評価する. TopicTiling[1]の既存手法である TextTiling[8]を提案した Hearst ら[8]の実験は, 被験者が判断した場面境界と TextTiling[8]の一致度を再現率, 適合率, F 値で評価を行っていた. そして, このときの F 値が 0.58 を超えていると, TextTiling[8]は十分に場面を分割できているという結果が得られていた. しかし, この方法は被験者の判断で結果が変わるため, 客観的に評価することが難しいと考えた. そのため, 本研究の評価は以下のように定義した式を使用し, 挿絵の挿入箇所と TopicTiling[1]の出力を比較する. また, 本実験で使用する TopicTiling[1]のトピックモデルは, トピックモデルの一般的な評価指標である Perplexity と Coherence で最良値が得られたモデルを用いる. 挿絵の挿入箇所とは, 青空文庫[15]のテキストにある [# 挿絵 n @.png 入る] (n: 文頭から数えた挿絵の番号, @: 任意のファイル名) を指す. 比較する手法として, 場面分割の既存手法である TextTiling[8]を用意した. TopicTiling[1]は文間で場面を分割する手法であるが, TextTiling[8]は単語間で場面を分割する手法である. そのままの TextTiling[8]を使用する場合, TopicTiling[1]と TextTiling[8]を比較することができないため, TextTiling[8]が文間で場面を分割できるように拡張した.

$$\text{再現率} = \frac{\text{場面境界との一致数}}{\text{挿絵の数}} \quad (2)$$

$$\text{適合率} = \frac{\text{場面境界との一致数}}{\text{TopicTiling[1]の場面境界の数}} \quad (3)$$

$$F\text{値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}} \quad (4)$$

3.2 重要文抽出の評価方法

重要文抽出の評価方法として一般的に ROUGE アルゴリズム[18]が使われている. ROUGE アルゴリズム[18]は ROUGE-N, ROUGE-L, ROUGE-S など複数の手法を持ち, 評価する重要文の集合と別の手法で用意した重要文の集合との類似度を求めて評価を行う. 今回の評価法は, N-gram (任意の N 文字が連続する文字列) の一致で評価する ROUGE-N と最長共通部分列の一致で評価する ROUGE-L を使用した. また, 重要文抽出に使用する実験データは Web サ

イト「小説家になろう」[19]の小説 100 冊である. 小説 100 冊は, 文の数が 1000 以上 5000 以下の条件を満たしているものをスクレイピングで収集した. 青空文庫[15]の小説で実験を行わなかった理由は, 青空文庫[15]にはテキストデータのみが用意されており, 小説の概要などの正解となる重要文が用意されていないからである. 一方, Web サイト「小説家になろう」[19]には, あらすじと本文が用意されている. しかし, このサイトの小説に挿絵は用意されていないため, 重要文抽出の実験のみ採用した. 本システムに採用した BERT[2]を用いる重要文抽出[3]は抽出型に分けられるため, 比較する手法は抽出型の既存手法であるグラフ構造のみで重要度を評価する LexRank[20]を用意した.

3.3 イラスト推薦の評価方法

小説の挿絵と解析結果に基づいて推薦したイラストの類似度を求めて評価を行う. 評価手法として, SSIM, MSE を使用した. このとき, 挿絵と本システムが出力するイラストのスタイルを統一するために画風変換を行う. また, 推薦した挿絵が文章解析後のテキストとどの程度類似しているのかを評価するためにコサイン類似度を使用した. 本研究は, オブジェクトが 1 つのイラストを推薦しているが, 青空文庫[15]に用意されている挿絵には複数のオブジェクトが存在している. そのため, 挿絵に複数のオブジェクトが含まれている場合は, OpenCV の findContours 関数で 1 つのオブジェクトを抽出した後に評価を行なった. また, Web サイト「いらすとや」[6]に用意されているイラストは単色の背景画像であるが, 青空文庫[15]の挿絵には複数の色が使用されている背景が存在する. 本実験では, 青空文庫[15]の挿絵の背景情報の影響を抑えるために, Canny 法[21]で輪廓を抽出して単色の画像に置き換えた.

4 実験結果

4.1 場面分割の実験結果

隣接する単語を含める範囲を決めるパラメータのウィンドウサイズを変更しながら TopicTiling[1]と TextTiling[8]を比較した結果を表 1 に示す. 表 1 より, TopicTiling[1]は既存手法の TextTiling[8]を上回る結果が得られた. また, 再現率, 適合率, F 値の最大値を得るウィンドウサイズは 10 であることが読み取れる. したがって, 場面を分割した後に行われる重要文抽出, 文章解析, イラスト推薦で扱うテキストは, ウィンドウサイズが 10 の TopicTiling[1]で場面を

分割したものである。TextTiling[8]が TopicTiling[1]に劣った理由を述べる。拡張した TextTiling[8]は文の類似度と閾値を比較して判定を行っていたが、各文の類似度がほぼ等しい場合適切に場面を分割することが難しくなる。各文の類似度がほぼ等しくなる理由として、文中に出現頻度が多い単語が周辺の単語に存在していれば、単語間の類似度に大きい変化がなくなり、文の類似度の平均にほぼ変化がないことが挙げられる。一方、TopicTiling[1]はトピックモデルを使用することで単語の出現に対して補正が行われ、類似度が大きく変化するため、適切に場面を分割することが可能になった。

表1 場面分割の実験結果.

手法(ウインドウ)	再現率	適合率	F 値
TopicTiling(5)	0.81	0.81	0.81
TopicTiling(10)	0.81	0.82	0.81
TopicTiling(15)	0.80	0.82	0.81
TextTiling(5)	0.46	0.77	0.50
TextTiling(10)	0.36	0.73	0.43
TextTiling(15)	0.40	0.72	0.45

4.2 重要文抽出の実験結果

既存手法のLexRank[20]と比較した実験結果を表2に示す。表2より、BERT[2]を用いる手法[3]がLexRank[20]を全て上回っている。すなわち、本システムの重要文抽出に採用した手法は、既存手法より有効であることが示される。LexRank[20]が劣った理由を述べる。LexRank[20]は、コサイン類似度でグラフを構築するため、各文の共通情報を持つ簡潔な文を抽出する可能性がある。この課題点が評価値を下げたと考えられる。本実験の結果は、尾崎ら[22]がニュース記事で実験を行なった結果に近い評価値が得られていることと、BERT[2]を用いる手法[3]は既存手法を上回ることは確認できた。本実験で使用したデータは小説のテキストであるため、全体の流れを考慮して重要文を判断する必要がある。しかし、本システムに使用している手法は、全体の流れを考慮して重要文を抽出するのではなく、学習済みモデルから重要文を予測して抽出する。解決する方法として、登場人物の関係や行動から場面の変化を高精度で予測できるようになることが挙げられる。

表2 「小説家になろう」[19]100冊の平均.

	ROUGE-1	ROUGE-2	ROUGE-L
BERT	0.33	0.06	0.18
LexRank	0.28	0.03	0.16

4.3 イラスト推薦の実験結果

実験結果を表3に示す。比較する分散表現としてDoc2Vec[12]を用意した。表3のSSIMより、BERT[2]とDoc2Vec[12]ともに評価値は0.72であった。このことから、青空文庫[15]の挿絵の情報を72%含んでいる。残りの28%の情報は、3つの問題点が影響していると考えられる。1つ目は、複雑なオブジェクトを含む挿絵から1つのオブジェクトを抽出する方法が、複数のオブジェクトを完全に削除できずに残った状態で出力されている可能性がある。その場合、適切に評価が行われなかったため評価値に影響を与える。2つ目は、イラスト内のオブジェクトの向きが考慮されていないことである。具体的には、挿絵のウサギが左を向いていることに対して、推薦しているウサギのイラストは正面を向いていると評価値が減少する。これを解決する方法として、オブジェクトの向きを統一して評価する方法が考えられる。この評価法の準備として、挿絵のオブジェクトの方向を自動的に判定する必要がある。しかし、イラスト内のオブジェクトの向きを自動的に判定する有効な方法は提案されていないため、今後の課題点となる。3つ目は、重要文抽出の影響を受けていると考えられる。4.2節の結果では、重要文抽出が有効に機能していないことが示されている。有効に機能しない場合、挿絵に必要な情報を抽出することは困難になる。そのため、挿絵に不必要な情報で推薦が行われた可能性がある。この課題点を解決するためには、小説のテキストを解析することで、場面の変化を読み取り、重要な文を適切に抽出する手法が必要になる。表3のMSEより、Doc2Vec[12]の方がBERT[2]より誤差が少ないことが読み取れる。すなわち、Doc2Vec[12]はBERT[2]より挿絵の色に近いイラストを推薦していると考えられる。最後に、コサイン類似度の実験結果について述べる。表3のコサイン類似度より、BERT[2]のコサイン類似度は0.79であるため、文章解析後のテキストの内容に対して高い関連度を持つ挿絵を推薦できていると考えられる。一方、Doc2Vec[12]のコサイン類似度は0.69であり、類似度は減少している。減少した理由として、Doc2Vec[12]の単語の内部状態が文脈によって変化しないことが挙げられる。すなわち、文脈を利用した推薦を行うにはBERT[2]の分散表現を用いることが有効である。

重要文抽出の影響を確認するために、重要文抽出を行わなかった結果を表4に示す。表3と表4より、イラスト同士を比較する評価値(SSIM, MSE)とコサイン類似度にはほぼ影響を与えていない。そのため、重要文抽出はイラスト推薦に対して十分に機能していないと考えられる。

表3 青空文庫[15]20冊の評価値の平均.

	SSIM	MSE	cosine
BERT	0.72	254.18	0.79
Doc2Vec	0.72	246.61	0.69

表4 重要文抽出を行わなかった時の結果.

	SSIM	MSE	cosine
BERT	0.72	254.28	0.78
Doc2Vec	0.72	247.66	0.69

5 おわりに

本研究では、場面分割、重要文抽出、文章解析、イラスト推薦の順で解析を行なった。BERT[2]で得られる分散表現に対してコサイン類似度を計算することで適切な挿絵を推薦するシステムを提案した。青空文庫[15]の小説20冊で実験を行なった結果、本システムは十分な精度で挿絵を推薦していることと、BERT[2]の分散表現を使用してコサイン類似度を計算することが有効に機能していることが示された。さらに精度を向上させるためには、重要文抽出を有効に機能させる必要がある。本実験では、BERT[2]の分散表現や Doc2Vec[12]の分散表現を用いて重要文を抽出していた。しかし、これらの分散表現は内容に最も関係している文を見つけることが困難であるという結果が得られた。人物の相関図などのテキスト情報以外の情報も機械に与えることで、この問題を解決できる可能性がある。そのため、テキスト以外の情報が重要文抽出にどの程度影響を与えるのかについて調査を進める。

参考文献

- [1] Martin Riedl, Chris Biemann, TopicTiling: A Text Segmentation Algorithm based on LDA, ACL 2012 Student Research Workshop, pp. 37-42, (2012)
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv preprint arXiv:1810.04805, (2018)
- [3] Yang Liu, Mirella Lapata, Text Summarization with Pretrained Encoders, EMNLP-IJCNLP, pp. 3730-3740, (2019)
- [4] 工藤 拓, 松本 裕治, チャンキングの段階適用による日本語係り受け解析, 情報処理学会論文誌, Vol. 43, No. 6, pp. 1834-1842, (2002)
- [5] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto, Applying Conditional Random Fields to Japanese Morphological Analysis, EMNLP, pp. 230-237, (2004)
- [6] いらすとや, <https://www.irasutoya.com/>, 最終閲覧日: 2020年11月5日
- [7] 西岡 渉, 沼尾 雅之, 小説を対象とした知的ナビ機能の付いた電子書籍リーダーの提案, DEIM Forum 2016 E8-5, (2016)
- [8] Marti A. Hearst, Multi-Paragraph Segmentation of Expository Text, the 32nd Annual Meeting of Association for Computational Linguistics, pp. 9-16, (1994)
- [9] 村井 聡一, 牛尼 剛聡, 電子化された小説の選別を支援する「立ち読み」インタフェース, 情報処理学会論文誌, Vol. 6, No. 4, pp. 24-34, (2013)
- [10] Karen Spärck Jones, A Statistical Interpretation of Term Specificity and its Application in Retrieval, Journal of Documentation, Vol. 28, No. 1, pp. 11-21, (1972)
- [11] 小坂 直輝, 小林 哲則, 林 良彦, 隠れた良作を推薦可能な Web 小説レコメンドシステムの提案, 人工知能学会インタラクティブ情報アクセスと可視化マイニング研究会(第23回)研究発表予稿集, pp. 1-8, (2019)
- [12] Quoc Le, T Mikolov, Distributed Representations of Sentences and Documents, International Conference on Machine Learning, pp. 1188-1196, (2014)
- [13] David M. Blei, Andrew Y. Ng, Michael I. Jordan, Latent Dirichlet Allocation, Journal of Machine Learning Research 3, pp. 993-1022, (2003)
- [14] 京都大学黒橋・村脇研究室, <http://nlp.ist.i.kyoto-u.ac.jp/index.php>, 最終閲覧日: 2020年11月5日
- [15] 青空文庫, <https://www.aozora.gr.jp/>, 最終閲覧日: 2020年11月5日
- [16] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, IEEE, Vol. 86, No.11, pp. 2278-2324, (1998)
- [17] 松本 裕治, 伝 康晴, 話し言葉の形態素解析, 自然言語処理研究会報告, Vol. 143, pp. 49-54, (2001)
- [18] Chin-Yew Lin, ROUGE: A Package for Automatic Evaluation of Summaries, Workshop on Text Summarization Branches Out, Post Conference Workshop of ACL 2004, pp. 74-81, (2004)
- [19] 小説家になろう, <https://syosetu.com/>, 最終閲覧日: 2020年11月5日
- [20] G. Erkan, D. Radev, LexRank: Graph-based Lexical Centrality as Saliency in Text Summarization, Journal of Artificial Intelligence Research 22, pp. 457-479, (2004)
- [21] John Canny, A Computational Approach to Edge Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679-698, (1986)
- [22] 尾崎 花奈, 小林 一郎, 事前学習モデルと潜在トピックを用いた文書要約への取り組み, 言語処理学会第26回年次大会発表論文集, p828-831, (2020)

非標準的に使用される単語の分散表現の補完手法

A method to complement distributed representation for non-standardly used words

魏 逸倫¹ 勞 瑛瑩 韓 東力²

Wei Yilun¹ Lao Yingying Han Dongli²

¹ 日本大学大学院 総合基礎科学研究科

¹ Graduate School of Integrated Basic Sciences, Nihon University

² 日本大学文理学部 情報科学科

² Department of Information Science, College of Humanities and Sciences, Nihon University

Abstract: 自然言語の分析においては、分析の障害となりうる隠語など非標準的に使用される単語が存在する。既存研究では未知語としての隠語に関する処理方法が多く述べられているが、別単語への置き換えとしての隠語を検出・分析する技術があまり開発されてこなかった。本研究では、トピックモデルと word2vec を用いて単語の分散表現を獲得し、入力文のトピックと文中に含まれた全単語の適合性を求めることで、隠語を自動的に検出し、そして、その正しい意味を補完する手法を提案する。

1. はじめに

自然言語の処理においては、単語の意味を分散表現として学習した上で利用するのがよくある手法であるが、学習された分散表現が処理対象となる単語にうまく対応できない場合が存在する。処理対象となる単語が分散表現の辞書に登録されていない、または登録された意味と異なる意味で処理されると、解析誤りが生じる。

これには、処理対象となる単語が分散表現を学習するために利用したコーパス内に出現しなかった「未知語」である場合と、単語が「誤字」や「隠語」などのように、人為的かどうかを問わず、辞書内で予想された形ではない、非標準的に使用された場合が存在する。例えば、次の2つの文に「パンダ」という単語を含んでいるが、文中では「パンダ」が動物という一般的な意味ではなく、「パトカー」という意味で使われている。すなわち、「パンダ」という単語が「隠語」として、本来の意味と異なる非標準的な使い方をしている。

- パンダに〇〇県警って書いてた。
- パンダの横っ腹に〇〇県警って書いてた。

非標準的に使用された単語の抽出及び分析に関する研究はいくつか存在しているが、概ねに分析目標となる単語を含むデータを大量に用いる手法[1][2][3]と、ある特定の領域におけるルールを用いる手法[4]、この2つの手法のいずれに属する。

ただし、「誤字」や「隠語」は性質上、同じ形式で表すことが少なく、大量のデータを用意するのが困難である。また、非標準的に使用された単語の出現ドメインや処理ルールを予想するのに限界があり、万遍なく対応することは難しい。そのため、既存の研究は高い精度が得られたものが多い反面、汎用性が低い。

本研究では、処理対象となる非標準的に使用される目標単語が以上の処理条件を満たさない場合でも利用できる、ルールベースに頼らないかつ目標単語を含む文章だけを必要とする、汎用性の高い手法の提案を目的とする。本手法により、処理される文書から目標単語を検出・分析することで、目標単語の元の意味となる単語を特定することを期待する。

2. 目標単語の抽出

本章では、目標単語の抽出について説明する。ただし、目標単語が分散表現の辞書に存在しない場合、抽出が簡単であるため省略する。隠語として文書に出現した単語が辞書に登録済みの場合は図1の通り、学習データを利用して単語の分散表現を求め、そして判別対象となる文に含まれている全単語の分散表現を用いて目標単語が含まれているかどうかを判断する。含まれると判断された場合に目標単語を抽出する。以下では提案手法を詳しく説明する。

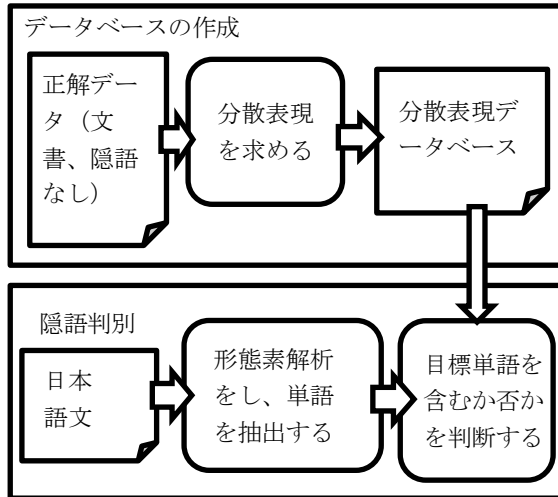


図 1 目標単語の抽出の流れ

2.1. 単語の分散表現

分散表現とは単語をベクトルに変換すること、もしくは変換したベクトルのことである。分散表現の学習方法として、word2vec とトピックモデルの一つである LDA を利用し、ベクトルの次元数およびトピック数を 200 と設定している。

LDA とは、文書中の単語がどのようなトピックを表しているかを確率的に求める言語モデルの一つである。この手法を利用すれば、文書に出現した全単語がそれぞれのトピックに属する確率がベクトルとして得られる。word2vec も単語のベクトルを得る手法であり、単語の周辺語を利用し、単語の意味をベクトルの各次元で解釈する。本研究では word2vec と LDA を比較した上で、word2vec で得られた単語ベクトルの各次元をトピックと見なす。

日本語としての正しさを考慮し、Wikipedia 日本語版のテキストを学習データとして用いて、単語をベクトル化する。このようにして得られた 870374 文書、合計 593,670 単語をベクトル化し、単語のベクトルの各次元の値を単語が各トピックとの関連度と考える。ただし、word2vec の結果でベクトルの次元の値がマイナスとなると、単語はトピックと反対する方向に意味を持つと考えられるので、トピックとの関連度を値の絶対値をとる。

2.2. 目標単語の判別

任意の 1 文に対して、少なくとも 1 つのトピックが含まれていると考えられる。また、文に含まれている全単語が文のトピックに寄与する。

文があるトピックに属すると思われるとき、トピックが文内の全単語と何らかの関連をもつと考えら

れる。すなわち、文に含まれるすべての単語がトピックとの関連度が高いほど、文全体がそのトピックに属していると判断するのが適切である。したがって、文があるトピックとの関連度 AP は各単語とトピックの関連度を用いて式(1)の通り相乗平均、あるいは式(2)の通りエントロピーを用いて評価する。

$$AP_{t,c} = (\sum_{w \in t} R_{w,c})/n \quad (1)$$

$$AP_{t,c} = -(\sum_{w \in t} R_{w,c} \log(R_{w,c})) \quad (2)$$

t: 文

c: トピック

$AP_{t,c}$: 文 t がトピック c に属する関連度

w: 単語

$R_{w,c}$: 単語 w とトピック c の関連度

n: 文に含まれる単語数

文は全トピックに対して最大となる関連度をもつトピックと最も関連しており、文中の全単語も同一トピックに属すと考える。さらに、文中の全単語が同じトピックに属す可能性を式(3)のように計算し、その結果を一貫性と定義する。一貫性は小さいほど、全単語が同一の文に含まれる可能性が小さくなり、すなわち文に存在すべきではない目標単語が存在する可能性が大きくなる。

$$S_t = \max(AP_{t,c}) (c \in C) \quad (3)$$

t: 文

c: トピック

C: 200 個のトピックの集合

S_t : 文 t の一貫性

図 1 の通り、分析対象となる各文に対し、形態素解析した上で、各単語の分散表現を用いて、一貫性を計算し、目標単語の存在を判別する。

2.3. 一貫性の有用性実験

2.3.1. 実験データ

Apple Store [5] の商品レビューから人手で隠語を含まない 200 文を収集し、ランダムに 100 文ずつに分け、データセット 1 とデータセット 2 とする。データセット 1 の各文に対し、ランダムに選定した 1 つの単語を同じ発音を持つ別の単語に置き換えることで作成された目標単語を含むデータセット 3 を作成する。ここでは、選定された単語に対し、Mecab の辞書を用いて読み仮名が同一となる単語候補の中からランダムに 1 つ選択する。最後に 3 つのデータセットに対し一貫性を求める。

表 1 抽出の結果実験

分散表現 の計算方法	文の関連度の計算方法	一貫性の平均			目標単語の認 識度
		データセット 1	データセット 2	データセット 3	
Word2vec	相乗平均	3.9706	3.9195	3.8501	79%
	エントロピー	3.6971	3.6930	3.6862	66%
	相乗平均×エントロピー	13.5855	13.4777	13.1148	82%
LDA モデル	相乗平均	0.0209	0.0208	0.0188	66%
	エントロピー	3.3363	3.3382	3.0817	61%
	相乗平均×エントロピー	0.0581	0.0582	0.0535	60%

2.3.2. 実験結果および考察

表 1 は実験結果のまとめである。文に目標単語を入れると、目標単語が文のトピックに属さないため、適切性が下がる。どの手法においても、目標単語を含むデータセット 3 がそのもととなるデータセット 1 より一貫性の平均値が小さくなっていることを確認できた。目標単語の認識度は、同じ文に対し、目標単語を含む方の一貫性が小さくなる割合を表しており、最大 82% で最小でも 60% といずれの手法を利用しても一定の有用性が確認できた。

また、異なる文を利用したデータセット 2 とデータセット 3 を比べても、目標単語を含むデータセット 3 の一貫性が小さく、一貫性は目標単語を含むかどうかを判断する一つの基準として利用できる可能性が示唆された。

さらに具体的な精度を求めるために、人手で選んだ目標単語を含む文 100 文と含まない 50 文を抽出しテストした。結果としてエントロピーを利用して文関連度を計算する場合に、精度が一番高く、 $f=0.9$ となる。ただし目標単語を判別する閾値として、表 1 に示された、データセット 2 とデータセット 3 の一貫性の中間値を利用した。

2.4. 目標単語の抽出

目標単語が含まれていると判定された文に対し、文内に含まれた全単語の平均ベクトルとの \cos 類似度が小さい単語を目標単語として抽出する。目標単語が 1 個だけ含まれると仮定し、類似度が最小となる単語だけを目標単語として抽出した場合、2.3.2 のデータに対して、0.69 の精度で目標単語を抽出できた。

3. 意味の補完

手法に汎用性を持たせるために、すべての単語がもっている、「意味」、「発音」、「形」の 3 つの要素を

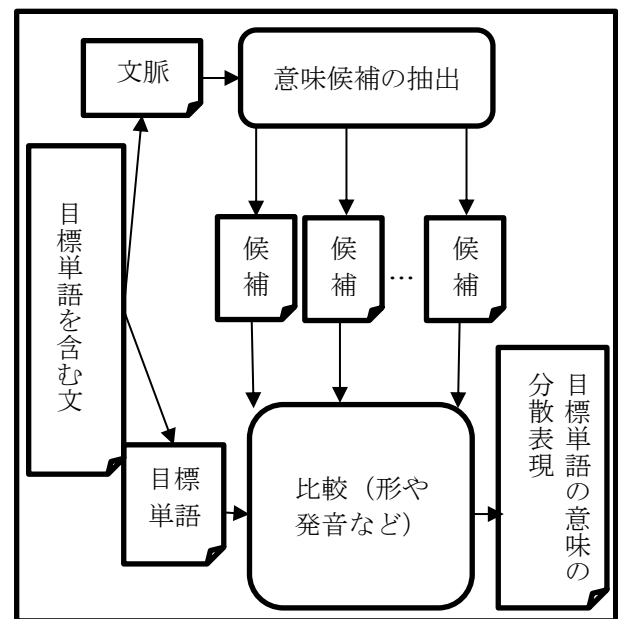


図 2 補完手法の流れ図

利用することで目標単語の意味を分析する。また、計算速度を上げるために、文脈から意味の候補を抽出してから目標単語の分析を行う。流れは図 2 の通りになる。

3.1. 文脈からの候補抽出

すべての日本語単語と比較して正解を探し出すことは現実的ではないため、文脈を利用することで、文章に関連する単語だけを取り出す。目標単語と関連を持つ可能性が大きい手がかりとして以下の 4 つを利用する。

- 目標単語を含む文書のタイトル
- 目標単語を含む文書全体
- 目標単語を含む文
- 目標単語の周辺語単語

文章に関連する単語は 4 つの手がかりから圧縮して得られたベクトルとの \cos 類似度が高い単語と定

義する。圧縮方法は図 3 の通り、ニューラルネットワークを利用して学習したものとなる。入力データは手がかりごとに含む全単語の平均ベクトルである。また本研究では分散表現ベクトルの長さを 200 と設定している。

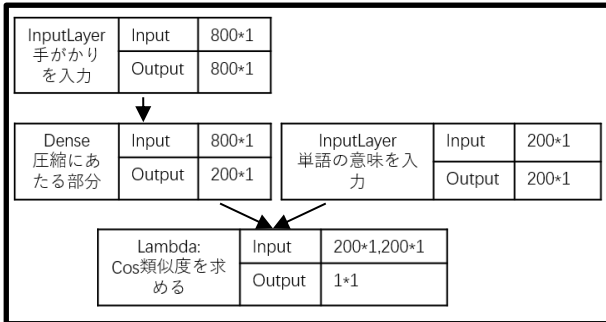


図 3 圧縮を学習するために利用した NN 構造

3.2. 候補の順番付け

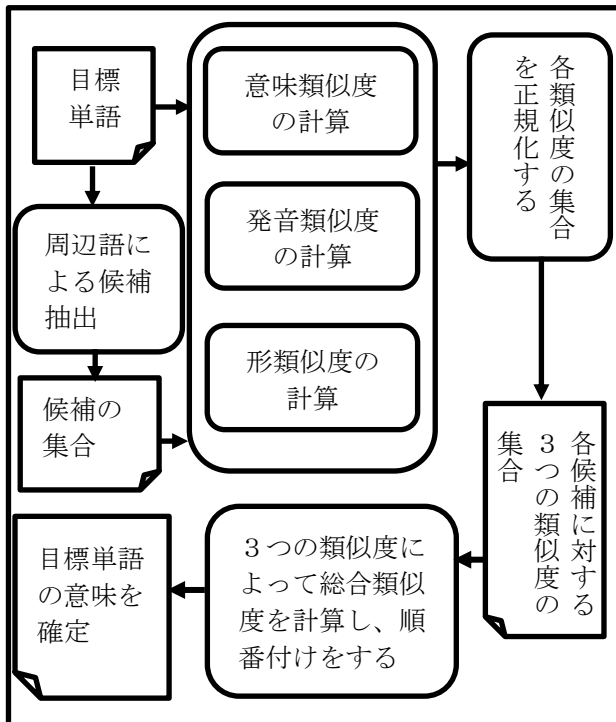


図 4 意味を確定する流れ

図 4 では候補を目標単語と比較することで、元の意味を確定する流れを示す。すべての候補に対し、目標単語と「意味」「発音」「形」で比較する。比較は以下のように類似度を求めることにより行う。

- 意味の類似度：分散表現ベクトルの cos 類似度
- 発音の類似度：単語ローマ字の編集距離
- 形の類似度：学習した CNN モデル

また、総合類似度は式(4)の通りに計算する。ただし、正規化により、類似度の値が 0 から 1 までの数字に変換する。

$$\text{総合類似度} = \text{MAX}(\text{意味類似度}, \text{発音類似度}, \text{形類似度}) \quad (4)$$

総合類似度が高いほど、候補と元の意味が一致する可能性が高い。

3.3. 事例による手法有用性の考察

ネット掲示板から抽出した目標単語を含む文を利用して、意味補完の効果を確かめた。元の意味は著者が文脈から判断し、決めたものとなる。

3.3.1. 事例 1

- 入力文: '一年間乙
 来年は分かりやすい相場だといいな〜

'
- タイトル: '〇〇〇〇年〇〇会を語る'
- 目標単語: 乙
- 分散表現の類似単語: '丙', '甲', '丑', '卯', '支路', '癸', '巳', '辰丸', '支文', '酉'
- 元の意味: お疲れ様
- 分析結果:

表 2 事例 1 の分析結果

順位	候補	文脈との意味類似度	発音類似度	形類似度	意味類似度
1	オウガ	0.9332	0.987	1	0.3932
2	オワタ	0.8169	1	1	1
3	イジリ	0.71065	0.3859	1	0.4007
4	イワオ	0.68545	0.987	1	0.2244
5	イチカ	0.64475	0.7048	1	0.2621
6	ルウム	0.41155	0.3859	1	0.6071
7	メラク	0.2024	0.7048	1	0.27045
8	綴る	0.7197	0.3859	0.2178	0.99995
9	負わさ	0.7032	0.99995	0.707	0.58795
10	緒形	0.5909	0.99995	0.1491	0.4882
11	わた	0.5143	0.99995	0.46555	0.68945
12	終わら	0.3541	0.99995	0.4621	0.6122
...

元の意味が 2079 番目にあり、全日本語の範囲内で考えると、上位であると言えるが、実用できる水準に至らなかった。

実際に上位となる単語は、総合類似度が高いが、「お疲れ様です」と比べると、文脈との関係性が低

いので、文脈を考慮する総合類似度の求め方の改善が課題となる。

3.3.2. 事例 2

- 入力文: 'オワタ<&喜びを表す顔文字>

'
- タイトル: '○○○○年○○会を語る'
- 目標単語: オワタ
- 分散表現の類似単語: '綴る', 'アドヴァタイジングスローガン', 'スラング', '悲しい', '失恋', 'キャプテン・ファルコン'
- 元の意味: おわった
- 分析結果:

表 3 事例 2 の分析結果

順位	候補	文脈との意味類似度	発音類似度	形類似度	意味類似度
1	載っ	0.8877	1	0.6741	0.7750
2	ボツ	0.7162	1	0.9765	0.7643
3	コツ	0.7099	1	0.5332	0.8115
4	ほっ	0.4915	1	0.8836	0.6521
5	モツ	0.4854	1	0.9437	0.9741
6	卒	0.3633	1	0.9481	0.9698
7	ボッ	0.3568	1	0.9943	0.6034
8	乙種	0.314	8	0.6764	1
9	コフィ	0.2921	3	1	0.3653
10	すっ	0.5217	2	0.9999	0.9094
11	蝦	0.3408	3	0.5934	0.9999
...
2079	お疲れ さま	0.8887	6	0.3356	0.5134
...

元の意味と近いと著者が判断した単語「終わら」が12番目に出現した。他の上位単語と比べると、周辺語との関係性が低く、自動的に判断するには難しいため、意味の補完は、図5のように、すべての上位単語をユーザーに見せ、人手で最終意味を決定するという半自動的な形式で行う予定となっている。

分析対象となる文章に出現した目標単語は、他の文書でも同じ意味で目標単語として使われる可能性が大きく、また同じ文書内で使われた他の目標単語と関連を持つ可能性が大きいため、既知の目標単語を新しい目標単語の元の意味を分析する手がかりと

して応用す手法も検討している。

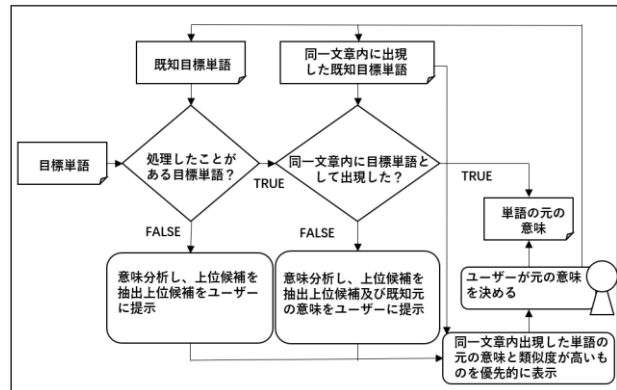


図 5 半自動的で意味を決める方式

4. まとめ

本研究では、自然言語処理の精度を上げるために、前処理として非標準的に使用される単語を検出し、その分散表現を補完する手法を提案した。

単語の分散表現を利用して、文のトピックと全単語の適合性を求めることで、隠語を自動的に検出する手法を提案した。評価実験では、約60%の精度で隠語が存在するかどうかを判断できることが判明し、手法の有用性が示された。

また、目標単語の意味分析において、目標単語の文脈と単語自身の特徴（形や発音など）を利用することで、意味を特定する手法を開発しているが、精度や文脈の利用などいくつかの課題が存在している。

今後の予定としては客観的な評価実験を実施し、本提案手法の有用性を確かめる。

参考文献

- [1] 大西 洋, 田島 敬史: 語の出現の偏りに基づく新たな目標単語の発見, 日本データベース学会論文誌, Vol. 12, No. 1, pp. 49-54, June 2013
- [2] 日田 仁: www サイト内の不正コンテンツ検出支援システムの構築, 分散システム/インターネット応用技術, 24-6 2001. 11. 30
- [3] 乾 亮, 山村 毅: 情報科学部視覚的「読み」を用いた分割表記文字の処理, 言語処理学会第25回年次大会発表論文集, P1-33, 2019
- [4] 屋 誠司: 常識的判断システムにおける未知語処理方式, 人工知能学会論文誌, 17, 667-675, 2002
- [5] <https://www.apple.com/app-store/>

あるある文章を用いたユーモア判定のための Word2Vec 学習モデルの検討

On a Recognition Model of the Sense of Humor "ARUARU" based on Word2Vec

野口啓太¹ 豊谷純² 大前佑斗²

Keita NOGUCHI¹, Jun TOYOTANI², Yuto OOMAE²

¹ 日本大学大学院 生産工学研究科 マネジメント工学専攻

¹Department of Industrial Engineering and Management, Graduate School of Industrial Technology, Nihon University

² 日本大学 生産工学部 マネジメント工学科

² Department of Industrial Engineering and Management, College of Industrial Technology, Nihon University

Abstract: Quantitative expression of humor is important for humans and computers to communicate organically in the future. One of the humor is the genre "ARUARU". There is a humorous story that people experience in their daily lives, and it is often talked about in everyday conversation. Therefore, it is effective for learning humor by machines. Hence, in this study, as the first step, we developed a model for humor recognition using word2vec method.

1. はじめに

Apple 社の Siri[1]や Amazon 社の Alexa[2]といった対話システムは、現代社会において広く浸透し始めている。さらに海外では不動産売買の場面や、体重体組成計に対話システムを活用する試みがなされている[3][4]。このことから、対話システムの活用は今後より広がると考えられる。一方、そのような社会において対話システムは、サービスや商品とユーザーをつなぎ続けるために、長期的かつ継続的に利用され続けることが求められる。その際、会話をしたい、または継続したいと思わせるデザインを、対話システムにもたせることは重要である。

上記のようなシステムを実現する上で、対話システムにユーモアの付与が有効であると考えられている。人同士の場面におけるユーモアの効果として、コミュニケーションの促進や社会的距離の縮小といったものなどがあげられる[5]。さらに、人と対話システムの場面においても、ユーモアは対話を継続させたい（以下、対話継続欲求）と思わせる要素の一つになるという報告もある[6]。このようなユーモアの効果から、実際の研究や試みとして、ユーザーからの任意の入力に対して関連するユーモア文章を自動生成するシステムや、なぞかけを自動生成するシステム、Twitter からユーモア文章を抽出し適切な文章で出力するシステムの検討といったものが行われている[7][8][9]。

先行研究においては、対話システムからユーザーに対してユーモアを発する試みが多くなされている。それらの試みは、ユーザーの対話継続欲求を満たす対話システムの実現において重要である。一方、日常会話において人々は他者などに自らユーモアを発することがある。これは、心理学においてユーモア表出という現象とされている。ユーモア表出の動機は、自分や他者を楽しませる目的や、自分や他者を支援する目的、他者を攻撃する目的などがあげられている[10]。確かに普段我々は友人や家族に対して会話の中で、自らユーモアを発し相手の反応を楽しんだりすることがある。また、対話継続欲求を高める要因の一つに“相手に話を聞いている実感を与える”というものがある[6]。したがって、システム側からのユーモアの出力に加えて、ユーザー側からのシステ

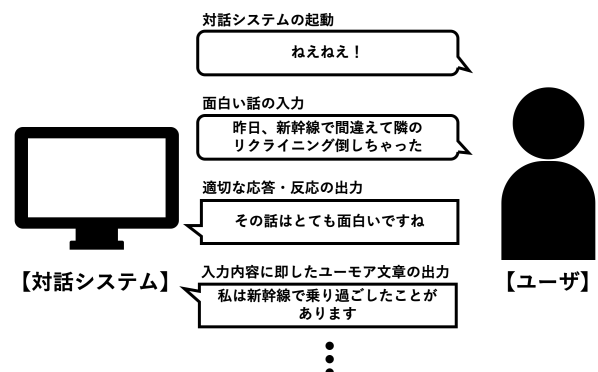


図1 本研究での理想的なインタラクションの流れ

ムに対するユーモア表出に適切に対応するようなデザインがあることで、ユーザの対話継続欲求がさらに高まる対話システムになると考えられる。

以上より本研究における、対話継続欲求を促すユーザと対話システムの理想的なインタラクションの流れは、図1のようなものになるとした。

一連の流れとして、始めにユーザが何らかの発話をするすることで、対話システムを起動する。次に、ユーザが自らの面白かったエピソードなどを話す。あくまでも日常会話の場面を想定しているため、ユーザのユーモア発話の内容は突拍子もないジョークやダジャレなどではなく、自然なエピソードトークとした。この部分に関して、人からシステムにそのような内容を語りかけるものなのか確証はない。しかし人は対話システムに対してもある程度、自身に関する情報を自ら話す傾向があることが明らかになっており、妥当な想定であると考えられる[11]。さらにその入力に対して、“面白いですね”や笑い声、さらに“ふーん”や“へー”といった入力内容に対して適した反応を出力する。その後、入力内容に関連したユーモア文章などを出力する。これにより自らの話を聞いてもらえている実感に加え、自身及び対話システムのユーモアの要素により、ユーザの対話継続欲求をさらに得られる対話システムになると考えられる。

このような対話システムを実現するためには、ユーザから入力される文章と、出力する文章のユーモアの度合いを判定する機構が必要である。そこで本研究では後述する Word2Vec を用いた、短いエピソードトークに対するユーモア度の判定手法の検討を目的とした。このユーモア判定の教師用データには、“あるある”を対象とすることとした。“あるある”とは「日常生活における事象・見聞などで、多くの人の共感を得ることのできる話題や、その話題で笑いをとる演芸のこと」(実用日本語表現辞典)とされている。その性質から、“あるある”は日常会話でのエピソードトークの内容と類似すると考えられるため、エピソードトークのユーモア度判定のための教師用データとして適していると言える。なお、本研究はユーモア判定の実現における第一段階の試みであるため、著者自身の感覚を再現することを目的とした。さらに判定は、面白いかつまらないかの二値分類とした。

2. “あるある”に関する考察

“あるある”とは多くの人々の共感を得る日常生活における事象や見聞とされている。つまり、多くの人々が経験したことのある、日常に関するエピソード

と捉えることができる。以下のような例があげられる。

- ・爪切りや耳搔きが行方不明になる
- ・会計した後にクーポンの存在に気づく
- ・スプーンを洗おうとすると水がかかる
- ・冷房の設定温度が定まらない
- ・電子マネーの残高に気づかず改札で止められる

一方で、人々が普段するエピソードトークとは、自身や自身の身の回りに起きた過去の成功や失敗、不運といった経験談にあたるものである。

これらから、普段人々がするエピソードトークの内容と“あるある”は類似したものになると考えられる。また図2に示す通り、普段の日常会話の場面でエピソードトークとして“あるある”を用いても、違和感はない。さらに、“あるある”は Twitter や Web サイトなどで容易に収集が可能である。よって、本研究の目的であるエピソードトークのユーモア判定において、教師用データとして“あるある”を活用することとした。なお、“あるある”は様々なジャンルが存在し、データを収集する時に際限がなくなってしまうため身近な話題である公共交通機関に関する“あるある”を対象にした。

3. Word2Vec について

本研究で用いる Word2Vec とは、Tomas Mikolov らによって提案された、自然言語処理に利用されるニューラルネットワークのモデルの一つである[12]。Word2Vec では、教師用データから単語の分散表現を作成し、単語同士のベクトル演算を可能にする。それにより単語と単語の類似度や、単語と単語の足し

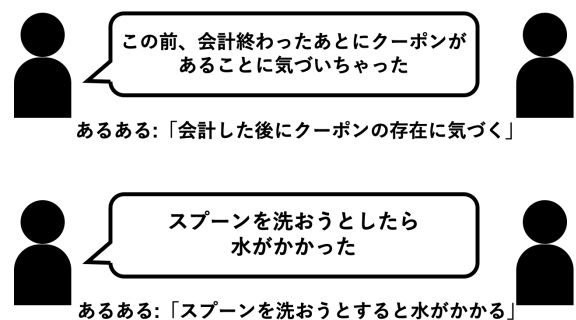


図2 エピソードトークを“あるある”にした場合

「新大阪/駅/から/電車/で/東京/駅/まで/行き/ます」
「新大阪/駅/から/東京/駅/まで/新幹線/で/通勤/する」

図3 分布仮説による類似度の例

「新大阪」/「駅」/「から」/「電車」/「で」/「東京」/「駅」/「まで」/「行き」/「ます」
 「新大阪」/「駅」/「から」/「東京」/「駅」/「まで」/「新幹線」/「で」/「通勤」/「する」
 「東京」/「駅」/「と」/「新大阪」/「駅」/「間」/「で」/「新幹線」/「が」/「遅延」/「する」
 「東京」/「駅」/「から」/「電車」/「と」/「バス」/「で」/「新大阪」/「駅」/「に」/「行く」

図4 類似度と共起度の関連

算や引き算が可能になる。これらの技術は、“単語の意味はその周辺語の出現頻度によって形成される”という分布仮説というアイデアに基づいている。仮説分布における単語同士の類似度のイメージ例を図3に示す。図3の例では、“電車”と“新幹線”という単語の、それぞれの周辺語には新大阪や東京、駅といった単語がある。つまり、この場合電車と新幹線という単語は、周辺語が似ているため類似度が高いものになると予測される。さらに、学習に用いた他の文章においても“電車”と“新幹線”の近くに新大阪や東京、駅といった周辺語が多く登場した場合、より類似度は高まる。

一方で、自然言語において共起という関係がある。共起とは文や文章の単位において、二つの単語が同時に出現することを示す。いくつかの文章で同じ単語同士が共起した場合、それらの文章の内容は類似したものになると考えられる。さらに、それらの文章で共起した単語の周辺語も類似したものになり、共起した単語同士の類似度も高まると考えられる。例として図4を示す。新大阪と東京の単語同士は共起関係にあり、さらに複数の文章で共起しているため共起度が高いと言える。また各文章で東京と新大阪が共起しているため、各文章は似た内容になっている。さらに、新大阪の周辺語には駅や電車、新幹線、東京といった単語が頻出し、東京の周辺語にも駅や電車、新幹線、新大阪といった単語が頻出しており、周辺語も類似している。つまり、図4の場合において東京と新大阪は類似度が高くなる。以上より、共起の度合いと類似度には関連があり、共起度が高い単語同士の類似度は高くなり、共起度が低い単語同士の類似度は低くなると言える。

4. 提案手法

本研究で用いたあるある文は、SNSのTwitter及びWebサイトにて収集した。また、あるある文のジャンルは公共交通機関での場面に限定した。さらに、ユーモア判定の学習における教師用データと検証用データとして、面白いあるあるとつまらないあるある文を、各40個ずつ用意した。また用意したあるある文は、表1のように教師用データとして面白いあ

るある28個とつまらないあるある28個に分け、検証用データとして面白いあるある12個とつまらないあるある12個に分けた。これらのあるある文の、面白いまたはつまらないの判断は、著者の感覚を再現するという本研究の目的のため、著者が行った。

4.1 Word2Vecの学習

何らかのあるある文を入力した際に、そのあるあるに対する評価をし、ユーモアを判定することが本研究の目的である。そのため、単語の分散表現を可能にするWord2Vecを用いて教師用データとした全56個のあるある文を対象に、あるある文のユーモア判定をするための学習済みモデルを作成した。以下にその手法を示す。

まず、学習に用いる全てのデータを図5に示す形式にした。この教師用データの形式は、あるある文に対して面白いか、つまらないかの評価を評価語として後ろに付け加えたものである。こうすることによって、評価語ごと単語ベクトルが学習され、あるある文における各単語と評価語のベクトル演算が可能になる。なお、学習前の分かち書きへの変換では、McCab[13]を使用し形態素解析を行なった。また、Word2Vecによる学習時のパラメータについての詳細を表2に示す。

4.2 あるある文のユーモア判定までの流れ

「新幹線の車内販売のアイスが硬すぎる」は「面白い」
 あるある文 評価語
 「終電は酔っ払いが多い」は「つまらない」
 あるある文 評価語

図5 教師用データの形式の例

表1 教師用データと検証用データの振り分け

	面白い	つまらない	合計
教師用データ	28	28	56
検証用データ	12	12	24
合計	40	40	80

表2 Word2Vecによる学習時のパラメータ

モデル	Skip Gram
文脈窓	5
単語最低出現回数	1
学習回数	3000
学習係数	0.1

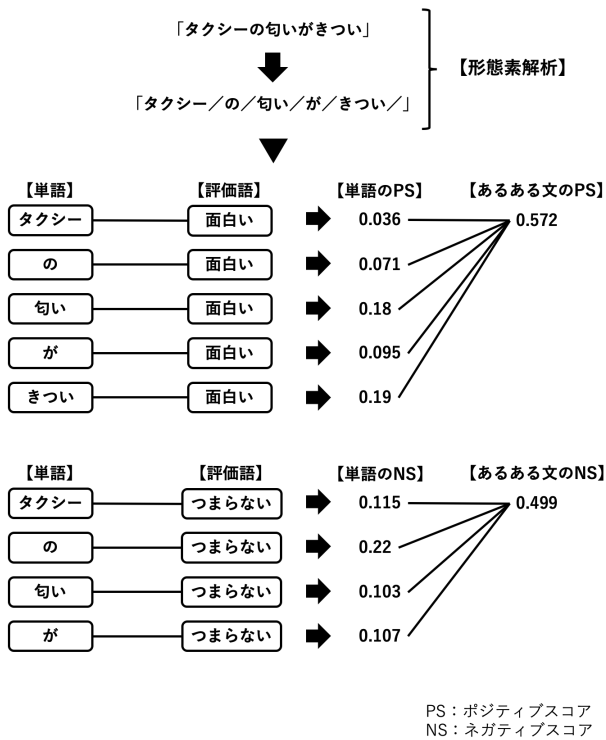


図6 あるある文のPSとNSの算出までの流れ

今回は、ユーモア判定の初期段階の試みであるため、あるある文に対するユーモア判定を“面白い”と“つまらない”の二値分類にすることとした。上記の学習済みモデルでは、評価語との共起度が高い単語は評価語との類似度が高くなり、評価語との共起度が低い単語は類似度が低くなる。つまり、本研究の場合“面白い”という評価語と類似度が高い単語は、面白いと評価されたあるある文に多く登場していることになる。さらにそういった単語が多く存在するあるある文は、面白いという判定をすることができる。そのため、各単語と“面白い”という評価語との類似度をポジティブスコアとし、“つまらない”という評価語との類似度をネガティブスコアとした。さらに、あるある文に対するユーモア判定の基準として、あるある文における各単語のポジティブスコアの合計とネガティブスコアの合計を比較し、スコアが大きい方の評価を採用することにした。図6にあるある文のポジティブスコアとネガティブスコアの算出までの流れを示す。図6の詳細として、入力された文章を形態素解析により分かち書きし、各単語を学習済みのWord2Vecに入れ、“面白い”や“つまらない”といった評価語との類似度を算出する。その後、各単語と“面白い”との類似度の総和を文章全体のポジティブスコアとし、各単語と“つまらない”との類似度の総和を文章全体のネガティブスコアとして算出する。

5. 結果

上記の学習済みモデルを用いて、検証用あるある文を対象に性能を検証した。検証用のあるあるには、面白いと評価したあるある文と、つまらないと評価したあるある文を各12個ずつ用いた。各検証用あるあるに対して、ユーモア判定した結果の例を表3に示す。さらに、二値分類の判定における混合行列を表4に示す。なお、混合行列におけるTP, TN, FP, FNは、それぞれ“TP:面白いあるあるを面白いと判定する”, “TN:つまらないあるあるをつまらないと判定する”, “FP:つまらないあるあるを面白いと判定する”, “FN:面白いあるあるをつまらないと判定する”という意味になる。また、表4における教師データとは、教師用データのあるある文を対象にユーモア判定をした際の結果であり、検証データとは検証用データのあるあるを対象にユーモア判定をした際の結果である。以上の結果より、本学習済みモデルの正答率は0.58であった。また、検証用データに対する適合率は0.57であり、再現率は0.67であった。この際の適合率とは、システムが面白いあるあると判定した全ての出力における、実際に面白いと評価されたあるあるの割合である。一方、再現率とは実際に面白いと評価されたあるあるを、システムが取り漏らさずに検出できた割合である。つまり、本モデルは、“つまらない”あるあるを“面白い”と判定してしまうこともあるが、実際に“面白い”あるあるの取り漏らしは少ないモデルと言える。

さらに、あるある文の入力からユーモア判定までの全体の流れのフローチャートを図7に示す。

表3 検証用データに対するユーモア判定の結果

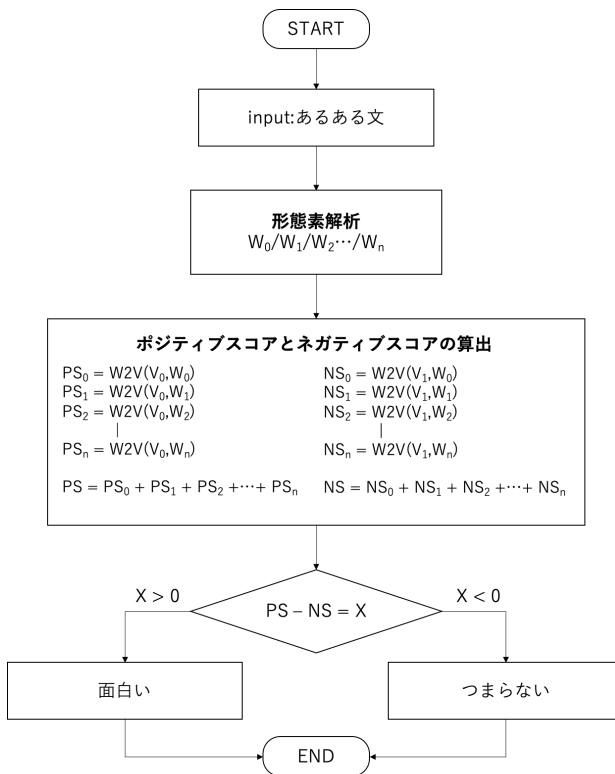
あるある文	PS	NS	実測値	推定値
飛行機で席を交換して欲しいと頼まれる	1.071	0.708	1	1
満員電車は降りる時が大変	1.003	0.644	1	1
電車で寝ようとする寝られない	0.977	0.704	1	1
電車で隣に座った人の肘が当たるのが気になる	0.977	1.356	1	0
バスは当たり前のように遅れる	0.377	0.860	1	0
飛行機で隣の席が外国人だと緊張する	1.066	1.289	0	0
観光地のタクシー運転手は観光客に優しい	0.685	1.084	0	0
雨の日はバスの湿気が高い	0.487	1.042	0	0
無意識に広告を見ている	0.838	0.340	0	1
電車の中で寝ると頭が痛くなる	0.788	0.631	0	1

0:つまらない, 1:面白い

表 4 ユーモア判定結果の混合行列

	実測値	推定値	教師データ	検証データ
TP	1	1	25	8
TN	0	0	21	6
FP	0	1	7	6
FN	1	0	3	4

0: つまらない, 1: 面白い



input(): ユーザからのあるある文の入力
 W_i : あるある文におけるi番目の単語
 PS_i : i番目の単語のポジティブスコア
 NS_i : i番目の単語のネガティブスコア
 V_0 : “面白い” (評価語)
 V_1 : “つまらない” (評価語)
 $W2V(V_0, W_i)$: “面白い”とi番目の単語の類似度
 $W2V(V_1, W_i)$: “つまらない”とi番目の単語の類似度

図 7 ユーモア判定までの流れ

6. 考察

検証データでの結果では、本モデルは適合率の値より再現率の値の方が大きいものになった。これは、“FN: 面白いのにつまらないと判定する”の値より“FP: つまらないのに面白いと判定する”の値の方が大きいためである。したがって、本モデルはあるある文に対して“面白い”という判定を出しやすい特徴があると言える。

この原因の一つとして、多くのあるある文に登場する“電車”という単語が影響していると考えられる。“電車”は、面白いと評価されたあるある文では14

回の登場であったのに対し、つまらないと評価されたあるある文では6回の登場であった。つまり、“電車”は“つまらない”という評価語より“面白い”という評価語との類似度が高くなり、“電車”が登場するあるある文はポジティブに偏る傾向が出てしまった。そのため、“電車”が入ったつまらない評価のあるある文も面白いと判定された可能性がある。しかし、今回の研究において、想定される本モデル使用場面は、ユーザがフランクにユーモラスなエピソードトークを対話システムに対して話すというものである。このような場面においては、ユーザのユーモア発言を厳しく判定するのではなく、なるべく発言の大部分に対して面白いと判定し、評価するような反応を返した方がユーザにとって親しみやすさを感じると考えられる。よって本モデルの特徴は上記の場面において適切であると言える。

一方で、検証データにおいて全体的な正答率がそこまで良いとは言えない。この一因として、教師用データの示す場面が均一ではなかったためであると考えられる。つまり、電車の場面におけるあるあるは面白いものが多く、バスやタクシーの場面におけるあるあるはつまらないものが多かった際、バスやタクシーの場面で面白いと評価されたものも“バス”や“タクシー”といった単語と“つまらない”という評価語の結びつきが強くなるため、つまらないと判定されることが考えられる。よって、この問題を解消するために、教師用データをバランスが取れたものになるように整形する必要がある。以下にデータのバランスを取るための項目を示す。

- 各場面におけるあるあるの数を統一する
 (例: 電車の場면을 20 個, バスの場면을 20 個など)
- 各場面における面白いあるあるの数とつまらないあるあるの数を統一する
 (例: 電車の場面における面白いあるあるを 10 個, つまらないあるあるを 10 個など)

今後はユーモア以外の感覚を判定するなど、当手法の活用の幅を広げつつ、判定の精度をさらに向上させる方法を模索していく。

参考文献

[1] Siri, <https://www.apple.com/jp/siri/>, (参照: 2020/11/02)
 [2] Alexa とできること, <https://www.amazon.co.jp/meet-alexa/b?node=5485773051>, (参照: 2020/11/02)
 [3] T. W. Bickmore, J. Cassell: Relational agents: a model and implementation of building user trust, In Proc. CHI, pp. 396-403, (2001)

- [4] T. W. Bickmore, R. W. Picard: Establishing and maintaining long-term human-computer relationships, ACM Transactions on Computer-Human Interaction, TOCHI, Vol.12 No.2, pp.293-327, (2005)
- [5] J. Holmes: Sharing a laugh: Pragmatic aspects of humor and gender in the workplace, Journal of Pragmatics, Vol. 38, pp.26-50, (2006)
- [6] 宮澤幸希, 常世徹, 榊井祐介, 松尾智信, 菊池英明: 音声対話システムにおける継続欲求の高いインタラクションの要因, 電子情報通信学会論文誌, Vol.J-95-A, No.1, pp.27-36, (2012)
- [7] 藤倉将平, 小川義人, 菊池英明: 非タスク指向対話システムにおけるユーモア応答生成手法, 人工知能学会全国大会論文集, Vol.29, pp.1-4, (2015)
- [8] 山本賢人, 杉本徹: 大規模コーパスを用いたなぞかけ作成システム, 情報処理学会第82回全国大会講演論文集, Vol.82, No.1, pp.419-420, (2020)
- [9] 宅和晃志, 吉川大弘, ジメネス フェリックス, 古橋武: 非タスク指向型対話システムにおけるあるあるツイートからの共感誘発型発話生成手法に関する検討, 日本知能情報ファジィ学会誌, Vol.30, No.5, pp.744-752, (2018)
- [10] 上野行良: ユーモア現象に関する諸研究とユーモアの分類化について, 社会心理学研究, Vol.7, pp.112-120, (1992)
- [11] R. Higashinaka, K. Imamura, T. Meguro, C. Miyazaki, N. Kobayashi, H. Sugiyama, T. Hirano, T. Makino, and Y. Matsuno: Towards an Open Domain Conversational System Fully Based on Natural Language Processing, in Processings of the 25th International Conference on Computational Linguistics, pp.928-939, (2014)
- [12] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean: Distributed representations of words and phrases and their compositionality, In NIPS, (2014).
- [13] MeCab (和布蕪) とは, <https://taku910.github.io/mecab/>, (参照: 2020/11/02)

Word2Vec と階層型ニューラルネットワークを活用した パワハラメールの自動判定

Automatic Recognition of Power Harassment Mail using Word2Vec and Multi Layered Neural Network

平山 健一郎¹ 豊谷 純¹ 大前 佑斗¹ 村田 大治²

Kenichiro Hirayama¹, Jun Toyotani¹, Yuto Omae¹, Daiji Murata²

¹日本大学生産工学部 ²テクバン株式会社

¹College of Industrial Technology, Nihon University, ²Techvan Co., Ltd.

Due to the spread of power harassment, the increase in retirees has become a social problem. Therefore, in this research, we developed an automatic judgment system for power harassment mail using Word2Vec and a neural network with the aim of preventing power harassment damage. As a result, it was found that some emails can be judged correctly. The results of this research will contribute to the prevention of power harassment damage on e-mail and SNS, and will contribute to the creation of a comfortable working society.

1. はじめに

近年、パワーハラスメント（以降、パワハラ）が社会問題となっている。厚生労働省が都道府県労働局等に設置した総合労働相談コーナーに寄せられる「いじめ・嫌がらせ」に関する相談は年々増加しており（図1参照）、その中でもパワハラに関する相談が最多となっている。このような状況を鑑み我が国は2019年5月に、改正労働施策総合推進法（通称：パワハラ防止法）を成立させた。大企業では2020年6月、中小企業では2022年4月から施行される予定である[2]。パワハラ防止法により、事業主は職場におけるパワーハラスメント防止のために、雇用管理上必要な措置を講じることが義務となった。これにより、企業はパワハラを早期に発見することが必要となる。

ここで、いくつか考えられるパワハラの手段の一つとして、メールによる精神的な攻撃が考えられる。したがって、メールからパワハラを検出することができれば、パワハラを早期に発見することが可能であると考えられる。メールは自然言語により記述されるため、自然言語処理の各種技術を組み合わせれば、パワハラメールの検出が可能である。

自然言語処理を利用した研究は、我が国で多数の実績がある。例えば、北村ら[3]は日本語を母国語としない人に向けて、日本語の難易度を調整するシステムを開発している。戸次ら[4]は助詞などの機能語を自動検出するシステムを構築している。村越[5]は学校教育現場における教師が作成した授業改善報告書のテキストデータを解析する方法について言及し

ている。

その他にも多数存在する自然言語処理の研究を踏襲し、メールテキストからパワハラか否かを検出する手法の開発が取り組まれている（例えば、FRONTEOが2019年にこのシステムを開発した[6]）。ただし、自然言語処理によるパワハラメールの検出については、未だ歴史が浅く、発展途上の状況と言っている。また、日本語の解析という特性上、他の研究と比べ、取り組む研究者や技術者が相対的に少ないとも言える。

以上の背景を鑑み本研究では、自然言語処理と機械学習を組み合わせることで、パワハラメールを自動検出する手法を開発する。本稿では、本手法における技術的要素及びテキストデータの処理過程について言及する。

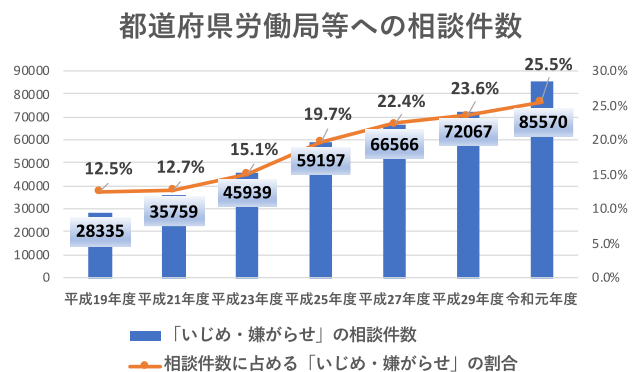


図1. 「いじめ・嫌がらせ」相談件数
(厚生労働省[1]の公開データより著者らが作成)

例：今日はいい天気です。

今日	は	いい	天気	です	。
・名詞 ・副詞可能	・助詞 ・係助詞	・形容詞 ・自立	・名詞 ・一般	・助動詞	・記号 ・句点

図 2. 形態素解析の一例

2. 要素技術と提案手法

本研究では、形態素解析、Word2Vec、ニューラルネットワーク（以降、NN）を組み合わせることでパワハラメールの自動検出する手法を開発している。本章ではこれらの概要について言及する。

2.1 要素技術：形態素解析と Word2Vec

形態素解析とは、与えられた自然言語を言葉の最小単位である形態素に分割し、それぞれの形態素の品詞情報を明らかにする技術である（図 2）。本研究では、Python 3.0 のインポートファイルである形態素解析ライブラリ MeCab と、辞書の一つである ChaSen を使用する。これを利用することで、「東京駅から電車で上野駅まで行きます」のような和文テキストを入力すると「東京/駅/から/電車/で/上野/駅/まで/行き/ます」のように、形態素に分解することができる。また各形態素には、名詞や助詞などの品詞情報も付与されている。日本語は英語などとは異なり、各品詞間がスペースで区切られることがないため、入力されたテキストのままでは、解析しにくい状態と言えるが、形態素解析を実施することで、この問題が解消されるという利点がある。

続いて、Word2Vec について言及する。これは、Mikolov et al.[7]らによって提案された、自然言語処理技術の一つである。構築された Word2Vec を利用することで、入力された単語と指定した単語との類似度を測定することができる。事前に与えられた膨大なテキスト内で、各形態素同士が同時に発生した程度に関する情報を参考とし、類似度測定を実現している。この技術を活用することで、数字として表現することができなかった形態素を、実数値のベクトルとして表現することが可能になる。

2.2 要素技術：ニューラルネットワーク

最後の要素技術として、階層型 NN を紹介する。これは、回帰あるいは分類問題を解くことのできる教師あり学習モデルの 1 つである。すなわち、入力と正答のペアで構成される教師データを用意することで、特定の入力に対して妥当な出力を推定する機能を構成することができる。入力層、中間層、出力層と呼ばれる三つのレイヤーを有している。入力層に入力されたベクトルが、ウェイトパラメータ、バ

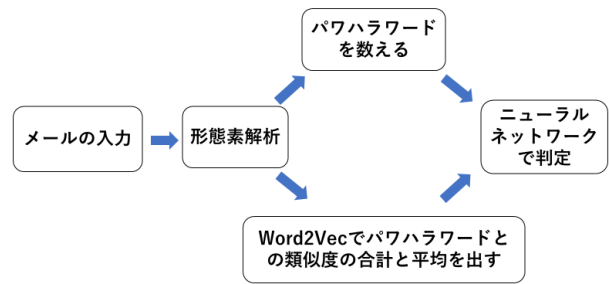


図 3. 提案手法の概要

イアスパラメータにより変換され、それが活性化関数に代入されることで、別の次元となった新たなベクトルが生成される。このベクトルを中間層と呼び、同様の演算により出力層に伝播する。出力層として算出された値が推定結果を意味する。

なお、NN の入力の実数値によるベクトルしか採用することができない。これは、自然言語であるテキストをそのまま入力することができないことを意味する。このため、テキストから何らかの状態を自動判定するために NN を利用する際は、Word2Vec などの技術を利用して、テキストを実数値に変換することが要求される。このように、入力されたテキストデータが、何らかの意味を持つ数値情報として変換されたものを、特徴量と呼ぶ。すなわち、特徴量が NN の入力層に与えられるベクトルを意味する。

2.3 提案手法：自動判定モデル

続いて、これまで紹介した要素技術を用いて考案したパワハラメールの自動判定モデルについて言及する（図 3）。提案手法においては、多数のメールテキストを利用して、事前に Word2Vec を学習しておくことが必要となる。これにより、パワハラメールで使用されやすい形態素同士の類似度を高いと出力するような Word2Vec を構築する（例えば、「このバカ、給料下げろ。休日も出勤してやれ」という文があった場合、「バカ」、「給料」、「休日」間の類似度が高くなる）。

続いて、与えられたメールテキストを形態素解析により、形態素に分解する。その後、事前に登録しておいたパワハラワード (PW) の登場頻度をカウントする。PW とは、「バカ」「アホ」「死」などのように、パワハラメールで使用されやすい形態素のことである。すなわち、直接的にパワハラに関連する用語が使用されている場合、それを個数として記録する意味を持つ。

上記処理により、パワハラで直接使用される用語が登場したときに、パワハラだと判定されやすくなる NN を構築することが可能だと考えられるが、PW

表 1. 特徴量ベクトル (PW が 2 つの場合)

次元数	特徴量
1 次元目	PW1 の使用数
2 次元目	PW2 の使用数
3 次元目	PW1 との類似度の平均値
4 次元目	PW2 との類似度の平均値
5 次元目	PW1 との類似度の合計値
6 次元目	PW2 との類似度の合計値

PWn: n 番目のパワハラワード

が直接使用されないパワハラメールの存在も、容易に想像することができる。このため、すべての形態素について、事前に構築した Word2Vec に代入し、PW との類似度を算出する。すなわち、ここで算出される類似度は、パワハラメールで登場しやすい形態素であれば高くなり、普通のメールのみで登場しやすい形態素であれば低くなるのが予期される。すべての形態素について得られた類似度について、その平均値および合計値を算出する。この情報を利用することで、PW が直接利用されていなくても、それと類似度の高い用語が使用されていれば、パワハラだと判定できる NN が構築できると考えられる。類似度の平均値と合計値の 2 つを採用している理由は、様々なメールの長さに対応するためである。

以上、一連の議論を踏まえた上で構成した特徴量ベクトルは表 1 の通りである。1, 2 次元目は PW の使用数, 3, 4 次元目は PW との平均値, 5, 6 次元目は合計値である。なおこれは、PW の数が 2 個 (例えば、「バカ」「アホ」など) の場合に限る次元数である。PW は 3 個以上登録しても良い。1 つの PW あたり 3 つの特徴量が生成されるので、採用する PW の数の 3 倍が、NN の入力層に代入される特徴量ベクトルの次元数となる。

この特徴量ベクトルとパワハラか否かの正答値がペアとなったデータが教師データとなる。これを利用して階層型 NN を学習させることで、パワハラメールの自動判定モデルが構築される。

3. 実験

3.1 目的と概要

提案手法の信頼性を検証することを目的に、実験を実施した。このため、はじめに教師データを構築した。Word2Vec の教師データとして、厚生労働省のホームページに掲載されていた過去のパワハラ判例から抜粋した PW が含まれた文章、および、それに類似する文章を複数作成し、それらをランダムに結合させ、500 文のメールテキストを生成した。NN の教師データとして、上記 500 文の中から 20 文をラ

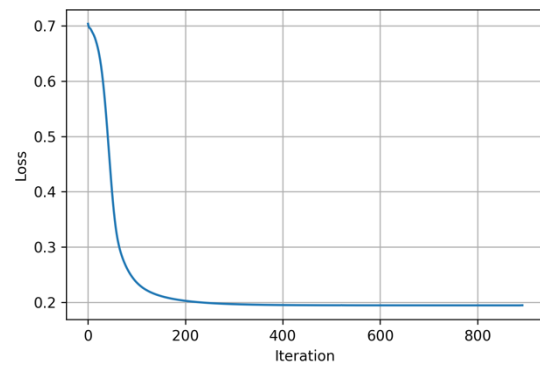


図 4. NN の学習過程

ンダムに選択したものをパワハラメール、著者らが作成した 20 文を普通のメール、合計 40 文のメールテキストを生成した。パワハラメールには 1、普通のメールには 0 とラベリングしている。

2.3 節で述べた PW については、「バカ」と「アホ」の 2 つを採用した。すなわち、特徴量ベクトルの次元数はその 3 倍である 6 次元となる。NN としては、入力層、中間層 1、中間層 2、出力層の 4 層構造とした。入力層は特徴量となる 6 次元、中間層 1 と 2 は 10 次元、出力層は 1 次元となる。出力層は最小 0 最大 1 を取る実数値であり、0.5 以上ならパワハラ、0.5 未満なら普通のメールと判定される。最適化手法は Adam、活性化関数はシグモイド関数、学習回数は 1000 回、学習係数は 0.01 をそれぞれ採用した。PW の類似度測定に利用する Word2Vec は、Skip-gram 法を採用した。登場回数が 1 回未満の単語は除去し、前後 5 単語までを対象とするよう設定した。学習回数は 3000 回、学習係数は 0.1 を指定した。

3.2 結果と考察

3.1 節の条件で構築した NN の学習過程を図 4 に示す。横軸は学習回数 (Iteration)、縦軸の Loss は NN のコスト関数であり、低くなるほど学習が進んでいることを意味する。図 4 を確認すると、学習回数が 100 回までで Loss が急激に減少し、200 回以降はほぼ収束していることが確認できる。学習回数が 1000 回に至る前、900 回付近で Loss が改善されなくなり、学習が打ち切られている。Loss が低いことから、教師データとして設定したメールについて、パワハラかどうか正しく判定できるようになったことが確認できる。

構築したモデルを利用して、いくつかのメールテキストにおいてパワハラか否かを自動判定した結果を表 2 に示す。ID 列の A は、「PW が直接使用されている短いパワハラメール」、B は「PW が直接使用されている長いパワハラメール」、C は「PW が使用

表 2. 提案手法による自動判定結果

ID	メール本文	字数	実測	推定	出力
A1	バカ。使えない。おまえ。おまえの存在が目障りだ。何時間かかっているんだ、俺だったら 10 分で出来るわ。	49	パワハラ	パワハラ	87.2 %
A2	いつでもおまえを飛ばすことができる。有休取れるわけないだろ。アホ。どれだけ迷惑かければ気が済むんだ。	50	パワハラ	パワハラ	99.5 %
A3	面接のときの発言はうそだったのか？だまされたわ。バカ。アホ。クソ。金返せ。もう契約更新しないから。	49	パワハラ	パワハラ	98.8 %
A4	意欲がない。アホ、そんなこともできないのかバカ。甘えるな。何時間かかっているんだ、俺だったら 10 分で出来るわ。	54	パワハラ	パワハラ	99.0 %
A5	おばさん。消えろ。反省してるのか。Fラン大学のバカ。すごい優秀な新人が来るぞ、おまえやばいな。	47	パワハラ	パワハラ	99.9 %
B1	契約とれるまで帰ってくるなバカ。殺す。すごい優秀な新人が来るぞ、おまえやばいな。いつでもおまえを飛ばすことができる。アホ。高卒。上司に逆らうのか？おっさん、ちゃんと考えたのか？社会人として常識だ。無礼なやつだな。	108	パワハラ	パワハラ	99.9 %
B2	ふざけるな。バカ。ぶさいく。おまえより成長が遅いやつは見たことがない。給与を返せ。おまえの子供がかわいそうだ。くさいわアホ。みんなやってくるぞ、おまえ以外できてるぞ。アホ。死ぬ気でやれ。	92	パワハラ	パワハラ	99.9 %
B3	ゆとり。上司に逆らうのか？おまえのせいで、みんなイライラしてるぞ。死ぬ気でやれアホ。休みの日、勉強したことを報告しろ。これ、常識だから。ゆとり。上司に逆らうのか？そんなこともできないのかバカ。	97	パワハラ	パワハラ	99.9 %
B4	消えろバカ。有休取れるわけないだろ、アホか？休めと思うな。おまえ何回目だ、何回教えたらわかるんだ。やる気あるのか？のろま。金返せ。そんなこともできないのか。甘えるなバカ。消え失せろアホ。	94	パワハラ	パワハラ	100.0 %
B5	まじめにやってくる？どれだけ迷惑かければ気が済むんだ。休みの日、勉強したことを報告しろ。アホ、そんなこともできないのかバカ。おまえ以外みんな出来てるぞ。チビ。中卒バカ。こんな成績で正社員なれないぞ。	98	パワハラ	パワハラ	99.9 %
C1	そんなこともできないのか。甘えるな。何時間かかっているんだ。すごい優秀な新人が来るぞ、おまえやばいな。	50	パワハラ	パワハラ	98.7 %
C2	反省してるのか。どれだけ迷惑かければ気が済むんだ。面接のときの発言はうそだったのか？消えろ。	46	パワハラ	パワハラ	69.1 %
C3	ふざけるな。有休取れるわけないだろ、おまえやばいな。ゆとり。これ、常識だから。	39	パワハラ	パワハラ	98.9 %
C4	いつでもおまえを飛ばすことができる。消えろ。休みの日、勉強したことを報告しろ。	39	パワハラ	パワハラ	87.2 %
C5	意欲がない。おまえより成長が遅いやつは見たことがない。おまえ何回目だ、何回教えたらわかるんだ。	47	パワハラ	普通	39.4 %
D1	金返せ。もう契約更新しないから。まじめにやってくる？どれだけ迷惑かければ気が済むんだ。おまえの子供がかわいそうだ。やる気あるのか？おまえ以外みんな出来てるぞ。中卒、チビ。ふざけるな。	90	パワハラ	パワハラ	99.8 %
D2	おまえ何回目だ、何回教えたらわかるんだ？そんなこともできないのか。甘えるな。休みの日、勉強したことを報告しろ。いつでもおまえを飛ばすことができる。これ、常識だから。甘えるな。くさいわ	91	パワハラ	パワハラ	99.4 %
D3	何時間かかっているんだ、俺だったら 10 分で出来るわ。おまえより成長が遅いやつは見たことがない。おまえ何回目だ、何回教えたらわかるんだ。まじめにやってくる？くさいわ、本当くさいわ。	87	パワハラ	普通	34.7 %
D4	クソ。金返せ。もう契約更新しないから。有休取れるわけないだろ、ゆとり。おまえ何回目だ、何回教えたらわかるんだ。上司に逆らうのか？こんな成績で正社員なれないぞ。	79	パワハラ	パワハラ	99.6 %
D5	使えないおまえ。面接のときの発言はうそだったのか？だまされたわ。おまえの存在が目障りだ。すごい優秀な新人が来るぞ、おまえやばいな。甘えるな、消え失せろ中卒。	79	パワハラ	パワハラ	90.5 %
E1	この度は、お忙しい中、貴重な時間を割いていただき誠にありがとうございました。何卒よろしくお願い致します	51	普通	普通	1.7 %
E2	この度詳しくお話を伺いことができ、大変参考になりました。本当にありがとうございました。	46	普通	普通	8.4 %
E3	ご理解感謝いたします。とても助かります。それでは表題の件、ご確認のほどよろしく願いいたします。	48	普通	普通	9.4 %
E4	その後の進捗は、いかがでしょうか？期限の方が今週金曜日と迫っているため、本日 15 時までにご送付をお願いできますでしょうか。	61	普通	普通	12.9 %
E5	また、本メールと行き違いでご連絡をいただいておりますら申し訳ありません。	37	普通	普通	2.5 %
F1	さて、この度お客さまからご返却いただきましたクレーム品の調査結果を下記の通りご報告させていただきます。当サンプルが必要となる会議が今週金曜日と迫っているため、本日 15 時までにご送付をお願いできますでしょうか。	104	普通	普通	10.6 %
F2	この度は、こちらの不手際で混乱を招いてしまいましたこと、心よりお詫び申し上げます。なお、正しい内容のメールを「打ち合わせについて」という件名で再送させていただきました。それでは、よろしく願い致します。	101	普通	普通	1.4 %
F3	表題の件ですが、引き続き発生経緯の調査をすすめ、今後発生頻度を踏まえて対策を検討していきます。この度は、こちらの不手際で混乱を招いてしまいましたこと、心よりお詫び申し上げます。	88	普通	普通	1.0 %
F4	先日はお時間の無い中ご対応いただきありがとうございました。早速ですが、ご提案いただいた件につきまして、提案資料を作成していただきたくお願いいたします。それでは、よろしくお願い致します。	92	普通	普通	1.7 %
F5	お世話になっております。その後の進捗は、いかがでしょうか。ご不明な点等ございましたらご連絡ください。ご参考までに過去の資料を添付させて頂きました。すみませんが、ご確認のほどよろしく願いいたします。	99	普通	普通	20.2 %

されていない短いパワハラメール」、Dは「PWが使用されていない長いパワハラメール」、Eは「パワハラではない短いメール」、Fは「パワハラではない長いメール」をそれぞれ意味する。すなわち、ID列A, B, C, Dをパワハラと判定すれば正答であり、E, Fを普通と判定すれば正答となる。また、A, BはPWが直接利用されているためNNにとって難易度の低い問題、C, DはPWが使用されていないパワハラメールであるため難易度の高い問題になると考えられる。

出力列はNNの実数値による出力値であり、0を0%、1を100%となるように表記している。

表2におけるA, Bを見るといずれも出力値が高いことがわかる。すなわち、PWである「バカ」と「アホ」のいずれかが直接使用されている場合、メールテキストの長さに依存せず、パワハラだと検出できることがわかる。これは、表1に示すように、PWの登場回数の特徴量としているためだと考えられる。

表2におけるC,DはPWが直接利用されていない場合である。この条件であっても、多くのパワーハラスメントを正しくパワーハラと判定することに成功している。これは、Word2Vecにより得られるPWとの類似度を特徴量としたことが理由であると考えられる。一方、C5とD3について、誤判定が認められた。該当のテキストには「成長」や「出来る」といった前向きなワードも含まれていることが原因かもしれない。この改善策として、Word2VecとNNの双方の教師データに、前向きなワードを含むパワーハラスメントを追加する必要があると考えられる。

表2におけるE,Fは普通のメールである。これらはいずれも適切に判定されていた。これは、テキスト内にパワーハラワードが直接使われていないことに加え、PWとの類似度の平均値と合計値が低いことが原因だと考えられる。

4. おわりに

本研究では、パワーハラスメントという重要な社会問題の1つを解決するために、メールにおけるパワーハラを自動判定するシステムを、形態素解析、Word2Vec、NNを用いて開発した。表2に示すように多様なメールにおいて適切な判定をできることが確認された。ただし、教師データや精度検証用のデータは少数であるため、今後はより多くのメールテキストを用意することが必要である。また、PWとして「バカ」と「アホ」の2つのみを採用したが、提案手法はその他にも様々なワードを採用できるので、今後は、PWの種別を増やすとともに、それが検出精度に与える影響を検証していく予定である。

参考文献

- [1] 厚生労働省：データで見るハラスメント, <https://www.no-harassment.mhlw.go.jp/foundation/statistics/>, 2020.10.30 閲覧.
- [2] 厚生労働省：職場におけるハラスメントの防止のために, https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/koyou_roudou/koyoukintou/seisaku06/index.html, 2020.10.30 閲覧.
- [3] 北村達也, 住田真一, 孝橋一希, 文難易度の調整を支援するシステム「かぶとエディタ」, 日本語教育方法研究会誌, vol.22, no.1, pp.14-15, 2015.
- [4] 戸次徳久, 仁科喜久子, 機能語の抽出ツールの開発, 日本語教育方法研究会誌, vol.9, no.2, pp.10-11, 2002.
- [5] 村越亮治, 統計解析ソフト R によるテキストマイニングの学校現場での応用可能性-授業改善のためのアクション・リサーチ報告書の分析を例に. 神奈川県立国際言語文化アカデミア紀要, vol.2, pp.165-175, 2013.
- [6] PR TIMES, FRONTEO, AI を活用した電子メール監査システム「Email Auditor 19」にハラスメント情報を学習させたナレッジベースを追加, <https://prtimes.jp/main/html/rd/p/000000267.000006776.html>, 2019, 2020.

11.12 閲覧.

- [7] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean: Distributed Representations of Words and Phrases and Their Compositionality, In NIPS, 2014.

GDMに基づくインタラクティブトピックモデリングの提案

Proposal for Interactive Topic Modeling based on GDM

小林 賢治¹ 高間 康史¹ 柴田 祐樹¹
Kenji Kobayashi¹ Yasufumi Takama¹ Hiroki Shibata¹

¹ 東京都立大学 システムデザイン学部

¹ Faculty of System Design, Tokyo Metropolitan University

Abstract: 本稿では、GDM (Geometric Dirichlet Means) に基づくインタラクティブトピックモデリングの手法について提案する。トピックモデルは多数の文書データに含まれるトピックの解析に用いられるが、教師なし学習であるため分析者の期待する結果が得られる保証はない。本稿では、幾何学的計算に基づく GDM を採用し、インタラクティブなモデル変更に必要なパラメータや操作を提案する。各操作の適用事例に基づき、操作意図をモデリング結果に反映できることを示す。

1 はじめに

本稿では、GDM(Geometric Dirichlet Means)[2] に基づくインタラクティブトピックモデリングの手法を提案する。トピックモデリングは、文書集合に存在するトピックを分析する手法として利用される、教師なし学習の一種である [1]。一つの文書には複数のトピックが含まれていることを仮定して、文書に含まれるトピックやトピックに関連する単語などを確率分布として表すことができる。しかしトピックモデルは教師なし学習であるため、分析結果がユーザの期待するものとは異なる場合がある。この問題を解決するため、人間参加型学習を用いたインタラクティブトピックモデリングが提案されている [6]。この手法は、トピックモデリングの学習過程に分析者の知見を介入させることで期待する結果に近づける半教師あり学習である。計算機による出力結果を分析者が確認し、フィードバックを与えて逐次的に修正していく。

トピックモデルとして LDA(Latent Dirichlet Allocation) を用いた既存研究 [6] を参考に、本稿では GDM アルゴリズムを用いたインタラクティブトピックモデリングを提案する。GDM は幾何学的視点からトピックを推定するトピックモデリングの手法である。LDA において、文書はトピックを基底とする潜在トピック座標単体上に射影される。この解釈に基づき、GDM では文書集合に重み付きクラスタリングを適用した結果に幾何学的補正を加えて基底ベクトルを推定する。

提案手法では初期クラスタリングの結果に基づき GDM モデルを作成する。その結果を分析者が確認し、期待とは異なる部分を変更する操作を適用する。操作結果に基づきクラスタリングおよび GDM を再実行し、モデルを更新する。既存研究 [6] ではモデル変更操作と

して、add word, remove word, change word order, remove document, merge topic, split topic, add to stop words を提案しているが、本稿では文書クラスタリングに基づく GDM の特性を活かし、add word, remove word, change word order, add document, remove document, merge topic, split topic, add to stop words の 8 種類を提案する。

本稿では Livedoor ニュースコーパス¹から取得した文書集合を用いて、提案する操作により想定する効果が得られるかを検証する。トピック数を指定して GDM を実行し、得られた初期モデルに対して提案する 8 種類の操作のうち紙面の都合により 3 種類のみについてそれぞれ適用した結果、意図通りにモデルが更新されることを示す。

2 関連研究

2.1 トピックモデリング

トピックモデリングは文書集合の潜在トピックを推定する確率モデルである。文書は複数の潜在的なトピックから確率的に生成されていると仮定して、文書に関連するトピックの生成確率を示すトピック分布、トピックに関連する単語の生成確率を示す単語分布を推定する。

トピックモデリングの代表的な手法として LDA[1] や GDM[2] が挙げられる。本節では本稿で利用する GDM について説明する。

¹<https://www.rondhuit.com/download.html>

2.2 GDM(Geometric Dirichlet Means)

GDM は文書集合のトピックを推定するトピックモデリングの一手法である．GDM で使用する変数を表 1, アルゴリズムを Algorithm1 に示す．

GDM は幾何学的計算に基づいて単語分布 $\phi = (\phi_{k_1}, \dots, \phi_{k_{N_K}})$ を推測する．トピック k は N_V 次元空間上のベクトル ϕ_k , 文書は各トピックを頂点とする単体上に生成される．GDM は主に以下の二つのステップから単語分布 ϕ を推測する．

- 文書集合に対し重み付きクラスタリングを実行しクラスタ中心 $\mu = (\mu_{k_1}, \dots, \mu_{k_{N_K}})$ を求める (Algorithm1, 2 行目)
- 幾何学的補正により μ から ϕ を求める (Algorithm1, 5 行目)

GDM ではクラスタ数 (N_K) を指定し, 重み付きクラスタリングにより文書集合をクラスタに分割する．本稿では重み付き k-means を利用する．各クラスタがトピックに対応する．次に, 求めた μ に対して式 (1) に示す幾何学的補正を加えて単語分布 ϕ を求める (Algorithm1, 5 行目)．

$$\phi_k = C + m_k(\mu_k - C), k \in K \quad (1)$$

$$C = \frac{1}{N_D} \sum_{d \in D} \bar{w}'_d \quad (2)$$

$$m_k = \frac{\max_{d \in D_k} \|C - \bar{w}'_d\|_2}{\|C - \mu_k\|_2}, k \in K \quad (3)$$

$$\bar{w}'_d = \frac{\bar{w}_d \circ \gamma_d}{\|\bar{w}_d \circ \gamma_d\|_1} \quad (4)$$

式 (4) の \circ はアダマール積を表す．ただし, 式 (1) により計算される単語分布で $\phi_{kv} < 0$ となる場合を考慮して $k \in K$ において式 (5) により ϕ_{kv} を更新する (Algorithm1, 7 行目)．

$$\phi_{kv} = \frac{R(\phi_{kv})}{\sum_{v' \in V} R(\phi_{kv'})}, R(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases} \quad (5)$$

Algorithm 1 Geometric Dirichlet Means

- 1: D の中心ベクトル C を求める
 - 2: クラスタ中心 $\mu_k (k \in K)$ を求める
 - 3: $m_k (k \in K)$ を求める
 - 4: **for** $k \in K$ **do**
 - 5: ϕ_k を求める
 - 6: **for** $v \in V$ **do**
 - 7: ϕ_{kv} を更新
-

表 1: GDM で使用する記号

記号	説明
$D = \{d_1, d_2, \dots, d_{N_D}\}$	文書集合
$K = \{k_1, k_2, \dots, k_{N_K}\}$	トピック集合
$V = \{v_1, v_2, \dots, v_{N_V}\}$	語彙集合
$D_k \subseteq D$	$k \in K$ に割り当てられた文書集合
w_{dn}	$d \in D$ の n 番目の単語
$\epsilon_d (> 0)$	$d \in D$ の重さ
$\bar{w}_{dv} (\geq 0)$	$d \in D$ 内の全単語に占める $v \in V$ の出現割合
$\gamma_{dv} (> 0)$	$d \in D$ における $v \in V$ の重さ
N_d	$d \in D$ の単語数
$\theta_d = (\theta_{dk_1}, \dots, \theta_{dk_{N_K}})$	$d \in D$ のトピック分布
$\phi_k = (\phi_{kv_1}, \dots, \phi_{kv_{N_V}})$	$k \in K$ の単語分布
$\theta_{dk} (\in [0, 1])$	$d \in D$ での $k \in K$ の生成確率
$\phi_{kv} (\in [0, 1])$	$k \in K$ での $v \in V$ の生成確率

2.3 人間参加型学習 (human in the loop)

人間参加型機械学習 (human in the loop)[3][4] は, 機械による学習に人間の知見をインタラクティブに介入させる学習方法で, 機械による学習結果の間違いや想定と異なる部分を人間が発見しフィードバックを与える半教師あり学習である．まず初めに, 機械は通常の機械学習 (分類モデルやクラスタリングなど) を行う．人間が学習結果を見て機械にフィードバックを与え, それに基づき機械は再学習をする．このプロセスを繰り返すことで, 分類・予測精度が向上したり, 人間が希望する結果に近づくことが期待できる．

人間参加型学習は人間の知見を学習に加えることができるため, 学習データが不足している場合や学習結果に高い精度が求められる場合に用いられる．

2.4 COP K-means

COP K-means[5] は通常の K-means に制約を加えるクラスタリング手法である．制約はデータ対に対するものとして与えられ, 以下の二種類が定義されている．

- Must-link 制約: 2 つのデータが同じクラスタに配置されなければならないことを示す制約
- Cannot-link 制約: 2 つのデータが同じクラスタに配置されてはならないことを示す制約

制約はデータをクラスタに割り当てる際に利用される。データに Must-link 制約が適用されている場合、対となるデータがすでにクラスタに割り当てられていれば同じクラスタに割り当てる。Cannot-link 制約が適用されている場合、対となるデータとは異なるクラスタに割り当てる。データは制約に違反しない限り最も近いクラスタに割り当てられ、制約に違反した場合はクラスタリングは失敗する。

3 GDMに基づくインタラクティブトピックモデリング

提案手法では、最初に通常のトピックモデリングを行う。その後、分析者が結果を見てフィードバックを機械に与えることで、トピックモデルを修正する。GDMでは重み付き K-means を用いるが、提案手法では分析者のフィードバックを考慮するために、GDMの再実行時には COP K-means を用いる。すなわち、モデル変更操作はクラスタリングにおいて考慮される。本節では、クラスタリングにおいて修正対象となるパラメータを説明した後、8種類のモデル変更操作について説明する。

3.1 修正対象パラメータ

提案手法では、以下のパラメータをユーザからのフィードバックに基づき修正する。

- γ_{dv} : 文書 d における単語 v の重さ
- ϵ_d : d の重さ
- Must-cluster 制約 : 指定したデータが指定したクラスタに割り当てられなければならないことを示す制約
- Cannot-cluster 制約 : 指定したデータが指定したクラスタに割り当てられてはならないことを示す制約

γ_{dv} は初期値を 1 とし、これを変更することで各文書における単語の重要度を調整する。 ϵ_d は初期値を 1 とし、文書がクラスタ中心に与える影響を調整する。Must-cluster 制約、Cannot-cluster 制約は COP K-means においてデータをクラスタに割り当てる際に利用する。データに Must-cluster 制約が適用されていた場合、指定されているクラスタにデータを割り当てる。Cannot-cluster 制約が適用されている場合、指定されているクラスタ以外の最も近いクラスタに割り当てる。

3.2 改良操作

本稿では 8 種類の改良操作を提案する。操作概要、操作手順、変更するパラメータについての説明を行う。

3.2.1 add word

add word の引数はトピック $k(\in K)$ と単語 $v(\in V)$ であり、 ϕ_{kv} が k において他の単語よりも高くなるように修正する操作である。実行手順を以下に示す。

1. k に割り当てられていない文書から v を含む $D_{\text{add}} \subset D$ を選択する
2. k と D_{add} に Must-cluster 制約を適用する
3. $\gamma_{dv}, d \in D_k$ を変更する

手順 2 において、分析者は v を含む各文書の内容を見て D_{add} を選択する。 k, v を指定して add word を適用した場合の Newton 法を用いた γ 更新式を式 (6) に示す。ただし、 $\gamma_{dv} = \gamma_1$ とし、 k で最も出現確率の高い単語を v_H とする。 (n) は更新過程において n ステップ目の値であることを表し、本稿では終了条件を $f(\gamma_1^{(n)}) < 10^{-8}$ とした。 A は加速係数である。

$$\gamma_1^{(n+1)} = \gamma_1^{(n)} - Af^{(n)} / \frac{\partial f^{(n)}}{\partial \gamma_1^{(n)}} \quad (6)$$

$$f^{(n)} = \mu_{kv}^{(n)} - \{\mu_{kv_H}^{(n)} + (\mu_{kv_H}^{(n)} - \mu_{kv_{H-1}}^{(n)})\} \quad (7)$$

$$\mu_{kv} = \frac{1}{\sum_{d \in D_k} \epsilon_d} \sum_{d \in D_k} \frac{\tilde{w}_{dv} \gamma_{dv}}{\|\tilde{w}_d \circ \gamma_d\|_1} \epsilon_d \quad (8)$$

3.2.2 remove word

remove word の引数はトピック $k(\in K)$ と単語 $v(\in V)$ であり、 ϕ_{kv} を 0 に近づける操作である。この操作では、 D_k 内の全文書 d について γ_{dv} を非常に小さな値に変更する。本稿ではこれを 10^{-6} とした。

3.2.3 change word order

change word order の引数はトピック $k(\in K)$ と単語 $v'_1, v'_2(\in V)$ であり、 k の関連単語リスト内で両単語の位置を入れ替える操作である。これは、 $\phi_{kv'_1}, \phi_{kv'_2}$ の大小関係を入れ替えることに相当し、 D_k 内の全文書 d について両単語に対応した重み γ_1, γ_2 の値をそれぞれ、 $i = 1, 2$ について式 (9) に基づき更新することで実現する。計算には 2 変数非線形連立方程式を解くために 1 次近似 Newton 法を使用する。 J はヤコビ行列 ($J_{ij} = \partial f_i / \partial \gamma_j$)、 δ はクロネッカーデルタを表す。本稿では終了条件を $|f_1| + |f_2| \leq 10^{-8}$ とした。

$$\gamma_i^{(n+1)} = \gamma_i^{(n)} - \sum_{j=1,2} (J^{-1})_{ij}^{(n)} f_j^{(n)} \quad (9)$$

$$f_i = \frac{1}{\sum_{d \in D_k} \epsilon_d} \sum_{d \in D_k} (\gamma_{dv'_i} \bar{w}_{dv'_i}) - \sum_{j=1,2} (1 - \delta_{ij}) \mu_{kv'_j} \quad (10)$$

3.2.4 add document

add document の引数はトピック $k (k \in K)$ と文書集合 $D_{\text{add}} (\subset D - D_k)$ であり, k に D_{add} の文書を関連付ける操作である. 既存研究 [6] には存在しない操作であるが, 文書が操作主体となる GDM の特徴を生かすために本稿で提案する. 実行手順を以下に示す.

1. k 以外に割り当てられた文書から追加する D_{add} を選ぶ
2. k と $d \in D_{\text{add}}$ に Must-cluster 制約を適用する
3. $\epsilon_d, d \in D_{\text{add}}$ を変更する

手順 1 において, D_{add} は k 以外に割り当てられた文書の内容を分析者が見ることによって選択する. 手順 3 において, ϵ_d の値は式 (11) により求める. ここで, k' は d が所属していたトピック, old, new は d を k' から k へ移動する前後の状態をそれぞれ意味する. $\text{Dis}(k, d)$ は k のクラスタ中心と d のユークリッド距離である. この更新により, $\text{Dis}^{\text{new}}(k, d) < \text{Dis}^{\text{old}}(k, d)$ となる.

$$\hat{\epsilon}_d = \frac{\eta_d}{1 - \eta_d} E_k^{\text{old}} \quad (11)$$

$$\eta_d = 1 - \frac{1}{2} \frac{\text{Dis}^{\text{old}}(k', d)}{\text{Dis}^{\text{old}}(k, d)} \quad (12)$$

$$E_k^{\text{old}} = \sum_{d \in D_k^{\text{old}}} \epsilon_d \quad (13)$$

D_{add} 内の d と k の関係を強めるために, ϵ_d は 1 以上の値とする必要があるが, モデル修正操作を何度も繰り返した際に大きくなりすぎること防ぐ必要がある. そのため, 式 (11) で求めた値を以下に従い修正する.

$$\epsilon_d = \begin{cases} \min(\hat{\epsilon}_d, \max(1, 0.9E_k^{\text{old}})) & (\hat{\epsilon}_d > 1) \\ 1 & (\hat{\epsilon}_d \leq 1) \end{cases}$$

add word, add document の 2 つの操作において, どちらも文書を追加するという操作を行うが, add word は指定した単語の重さを変更することが目的であるのに対し, add document は文書をクラスタに近づける事が目的である点で異なる.

3.2.5 remove document

remove document の引数はトピック k と文書集合 $D_{\text{remove}} (\subset D_k)$ であり, k から D_{remove} の文書を削除する操作である. 実行手順を以下に示す.

1. D_k から D_{remove} を選ぶ
2. k と $d \in D_{\text{remove}}$ に Cannot-cluster を適用する
3. $d \in D_k - D_{\text{remove}}$ について $\hat{\epsilon}_d = \frac{1}{2} E_k^{\text{old}}$

手順 1 において, 分析者は k の各文書の内容を見て D_{remove} を選択する. 手順 3 では D_{remove} の各文書が k のクラスタ中心から離れるように, ϵ_d を求める. この更新により, $\text{Dis}^{\text{new}}(k, d) < \text{Dis}^{\text{old}}(k, d)$ となる. また, $\hat{\epsilon}_d$ において 3.2.4 節と同様の修正を行う.

3.2.6 split topic

split topic の引数はトピック k と単語集合 $V_{\text{add}} \subseteq V$ であり, k を二つのトピックに分ける操作である. 実行手順を以下に示す.

1. D_k において $v \in V_{\text{add}}$ を含む文書から, 新しいトピック k_{new} に移す D_{add} を選択する
2. D_{add} の重心を中心とするクラスタ k_{new} を生成する
3. $d \in D_{\text{add}}$ と k_{new} に Must-cluster 制約を適用する

手順 1 において, D_{add} は $v \in V_{\text{add}}$ を含む文書の内容を分析者が見ることによって選択する.

3.2.7 merge topic

merge topic の引数はトピック k, k' であり, 二つのトピックを一つに統合する操作である. 実行手順を以下に示す.

1. $d \in D_k \cup D_{k'}$ と k に Must-cluster 制約を適用する
2. K から k' を削除する

3.2.8 add to stop words

単語分布に出現する単語の中には分析者から見て不要な単語が含まれていることがある. add to stop words の引数は $v (v \in V)$ であり, v をストップワードに追加する操作である. ストップワードに追加された単語は GDM の分析対象外となる. $\bar{w}, \gamma, \mu, \phi$ から v に関する値を削除する.

4 予備実験

4.1 実験概要

予備実験では, 3.2 節で提案した操作によって GDM のモデルが意図通りに変更されることを検証する. 紙面の都合により, add word, add document, change word order の適用結果のみ示す. 学習データは Livedoor ニュースコーパス²を使用した. 前処理として Mecab を利用した形態素解析を行い, 形態素として名詞(一般, 固有名詞, サ変接続, 形容動詞語幹)と, 自立形容詞を抽出した. 抽出された形態素から, SlothLib³のストップワードリストに記載された単語を除去した. 文書数は 100, 語彙数は 2,402, 総単語数は 13,907 となった. 本稿では, モデル変更前後の評価指標として式 (14) で定義される Perplexity (P) を利用する. これは確率モデルの予測性能を評価する指標であり, 値が低いほど性能が高いことを示す.

$$P = \exp \left(- \frac{1}{\sum_D N_d} \sum_D \sum_{n=1}^{N_d} \log p(w_{dn}) \right) \quad (14)$$

$$p(w_{dn}) = \sum_{k \in K} \theta_{dk} \phi_{kw_{dn}} \quad (15)$$

4.2 実験結果

GDM による初期モデルの各トピックにおける文書割り当て状態を表 2 に, add word(3, 'mac') を適用した後の文書割り当て状態を表 3 にそれぞれ示す. この操作により, 文書 0, 1, 7 がトピック 4 から 3 へ移動している. また, 操作前後でのトピック 3 における単語分布 (ϕ_{k_3v} の上位 10 単語) を表 4 に示す. 表 4 より, 単語 'mac' のトピック 3 における生成確率が操作後に最も高くなっていることがわかる.

初期モデルに対し, add document(2, D_{add}), $D_{add} = \{81, 85, 86, \dots, 98\}$ を適用した結果を表 5 に示す. 表 5 より, 指定した文書が全てトピック 2 に割り当てられていることがわかる.

初期モデルに対し, change word order(3, 'ノート', '書籍') を適用した後のトピック 3 における単語分布 (ϕ_{k_3v} の上位 10 単語) を表 6 に示す. 表 6 より, 単語 'ノート', '書籍' のトピック 3 での生成確率における大小関係が入れ替わっているのがわかる.

操作前後の Perplexity を表 7 に示す. 表 7 より, 操作前と比べて add word, add document によって Perplexity は減少しており, モデルが改善されていることがわかる.

表 2: 初期モデルの文書割り当て状態

トピック id	文書 id
0	9, 17
1	62, 68, 70, 72
2	97
3	2, 8, 16
4	0, 1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 63, 64, 65, 66, 67, 69, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 99

表 3: add word(3, 'mac') 適用後の文書割り当て状態 (変化のあったトピックのみ掲載)

トピック id	文書 id
3	0, 1, 2, 7, 8, 16
4	3, 4, 5, 6, 10, 11, 12, 13, 14, 15, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 63, 64, 65, 66, 67, 69, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98, 99

表 4: add word(3, 'mac') 適用による単語分布の変化

初期モデル		add word	
単語	生成確率	単語	生成確率
'サービス'	0.0331	'mac'	0.0226
'ノート'	0.0228	'サービス'	0.0170
'生産'	0.0201	'セキュリティ'	0.0170
'ultrabook'	0.0192	'リリース'	0.0165
'レッツ'	0.0190	'os'	0.0145
'書籍'	0.0183	'lion'	0.0134
'電子'	0.0183	'ノート'	0.0133
'mac'	0.0174	'アップル'	0.0129
'hp'	0.0173	'製品'	0.0125
'終了'	0.0170	'パソコン'	0.0124

²<https://www.rondhuit.com/download.html>

³<http://svn.sourceforge.jp/svnroot/slothlib/CSharp/Version1/SlothLib/NLP/Filter/StopWord/word/Japanese.txt>

5 まとめ

本稿では、GDMに基づくインタラクティブトピックモデリングを提案した。提案手法では、分析者の意図をモデルに反映する8種類の操作を定義し、パラメータ更新方法を示した。ニュース記事を対象とした予備実験により、add word, add documentにより意図通りにモデルが変化し期待したトピックが得られることを示した。スペースの都合上示すことができなかったが、他の操作でも意図通りのモデル変更が可能であることを確認している。今後は、より多様な条件での検証を行う他、実際のユーザによる評価実験を行う予定である。

参考文献

- [1] D. M. Blei, A. Y. Ng, and M. I. Jordan: Latent Dirichlet allocation, *Journal of Machine Learning Research* 3, pp. 993-1022, 2003.
- [2] M. Yurochkin, and X. Nguyen: Geometric Dirichlet Means algorithm for topic inference, *Proceedings of the 30th International Conference on Neural Information Processing Systems*, pp. 2513-2521, 2016
- [3] H. Lee, J. Kihm, J. Choo, J. Stasko, and H. Park: iVisClustering: An Interactive Visual Document Clustering via Topic Modeling, *Eurographic Conference on Visualization*, vol. 31, no. 3, pp. 1155-1164, 2012
- [4] J. Choo, C. Lee, C. K. Reddy, and H. Park: UTOPIAN: User-Driven Topic Modeling on Interactive Nonnegative Matrix Factorization, *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, 2013
- [5] K. Wagstaff, C. Cardie, S. Rogers, and S. Schroedl: Constrained K-means Clustering with Background Knowledge, *18th International Conference on Machine Learning*, pp. 577-584, 2001
- [6] A. Smith, V. Kumar, J. Boyd-Graber, K. Seppi, and L. Findlater: Closing the Loop: User-Centered Design and Evaluation of a Human-in-the-Loop Topic Modeling System, *Proceedings of the 23rd International Conference on Intelligent User Interfaces*, pp. 293-304, 2018

表 5: add document 後の文書割り当て状態

トピック id	文書 id
0	9, 17
1	62, 68, 70, 72
2	81, 85, 86, 87, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98
3	2, 8, 16
4	0, 1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 63, 64, 65, 66, 67, 69, 71, 73, 74, 75, 76, 77, 78, 79, 80, 82, 83, 84, 88, 99

表 6: change word order(3, ‘ノート’, ‘書籍’) 適用による単語分布の変化

初期モデル		change word order	
単語	生成確率	単語	生成確率
‘サービス’	0.0331	‘サービス’	0.0328
‘ノート’	0.0228	‘書籍’	0.0225
‘生産’	0.0201	‘生産’	0.0202
‘ultrabook’	0.0192	‘ultrabook’	0.0192
‘レッツ’	0.0190	‘レッツ’	0.0191
‘書籍’	0.0183	‘ノート’	0.0186
‘電子’	0.0183	‘電子’	0.0182
‘mac’	0.0174	‘mac’	0.0173
‘hp’	0.0173	‘hp’	0.0173
‘終了’	0.0170	‘終了’	0.0169

表 7: Perplexity(P) の比較

操作	P
操作前	2645
add word	2538
add document	2286
change word order	2645

クラシックギターの指導現場における 知識の体系化とその実現に向けた検討

A Study of Classical Guitar Lessons for Knowledge Systematization

三浦寛也^{1*} 飯野なみ^{1,2}
Hiroya Miura¹ Nami Iino^{1,2}

¹ 理化学研究所革新知能統合研究センター

¹ RIKEN Center for Advanced Intelligence Project

² 産業技術総合研究所

² National Institute of Advanced Industrial Science and Technology

Abstract: This paper presents the features of the one-on-one classical guitar lessons of a particular pair. The aim of our research are clarify the patterns of interactions between speeches and performances and visualize the change of student's performances. In this study, we report the results of the acoustic analysis of the audio data and the conversation analysis of actual contents of speeches.

1 はじめに

本研究の目的は、演奏指導現場における音声データを用いた知識構造化データの獲得と楽譜へのアノテーションに基づく知識発見を支援するシステムの開発である。近年、音楽教育の場において、より細かな指導が重視される個人レッスンの需要が高まっている。当該場面においては、演奏に加え議論を交わすことがあるため、演奏練習の音声を収録することが多く、記録したデータを有効活用するためには、練習時の発話や演奏状況の効果的な記述が必要となる。このような背景から、筆者らは、これまで楽器演奏における教授学習場面を対象とした、より持続的に知識を蓄積し獲得していくためのナレッジマネジメントの基盤構築のための研究を進めてきた。具体的な取り組みとして、これまでおよそ5年に渡り、個人レッスンに着目してクラシックギターの個人レッスン内で起こる演奏や発話などの音声データを収録している。

楽器演奏の個人レッスンには、極めて個人差が大きい内容を孕んでいる。しかし共通して生徒の上達支援と指導の質的向上を目的としている。積み上げられていった

その過程から、学習者や指導者の思考を知ることは不可欠である。このような個別の事象を扱う研究では、ボトムアップ的に複数の事象の中から、共通となる特徴を発見していくことが期待できる。そこで本稿では、先述した目標達成のための第一歩として、特定のペアにおけるクラシックギターレッスンの複数回にわたる音声データを対象に、教授学習場面での発話と演奏のインタラクションのパターンや、実際の発話内容に着目した分析を進め、指導内容の変遷やその特徴、そして協働的な演奏構築過程がいかに行われるのかを明らかにする。

楽器指導に関する先行研究はこれまで教育学や芸術学の観点から述べられることが多かったが、近年では、コンピュータを用いた演奏練習の支援が数多く行われてきている [坂本 15]。具体的な事例として、遠隔隔で同期式のアンサンブル練習を行うためのプロトタイプツールキット [Akoumianakis 08] や、音楽初心者向けのリアルタイムな音楽コラボレーションシステム [Gurevich 06] などが開発されてきた。また、音楽理解や音楽学に関する支援のための研究では、演奏を録音し音程の情報を多角的に表示することでピッチの揺れやずれを指摘するシステム [Lim 09] や、オーケストラスコアの理解を支援するシステム [松原 12]、共同演奏における演奏者間のコミュニケーションモデルを調査により構築した研究 [Kawase 07] などがある。しかしこれらの研究の多くは、大型ディスプレイや専用アプリケーションなどの環

* 連絡先: 理化学研究所 革新知能統合研究センター, 〒103-0027
東京都中央区日本橋1丁目4-1 日本橋一丁目三井ビルディング 15 階, E-mail: hiroya.miura@riken.jp

境に関する制約が多く、実際の指導現場を科学的な観点から分析・支援する研究はあまりない。

2 本研究の最終目標

本研究の最終目標として、演奏指導現場におけるマルチメディア情報を用いた情報構造化と楽譜へのアノテーションによる、再利用可能な知識の抽出を実現するシステムの開発を実現する。この目標を達成するためには、楽器演奏の指導現場で記録された発話や演奏状況を効果的に記述するための枠組みを提案し、以下の機能を包含するようなシステムの開発が求められる；レッスン中に記録された発話を抽出し譜面上に表示する機能、各発話と楽曲との対応関係を特定する機能、タブレット端末等で撮影した写真や動画を読み込み、楽譜上の適切な箇所に関連付ける機能、発話情報とオントロジーとの関連付けをおこなう機能(図1)。

目標達成のための具体的なアプローチとして、指導者と学習者の発話内容やそタイミング、演奏音と発話との前後関係から、演奏指導に起こる発話がどの演奏音に対するものかを自動推定することで、構造化された知識データを生成し、そこから再利用可能な知識の体系化を目指す。また、レッスンで起こる発話には、同一あるいは類似テーマという関係付けを同定し、類似性計算やアノテーションの付与によって、レッスン集合を構造的に表現することができると考えている。本研究では、その初期検討として、主に音声区間分析および発話分析について取り組み、そこで得られた成果や知見をもとにシステムの構築を目指す。

3 教授学習場面の音声区間分析

本節では、レッスン中に起こる事象を正確に記述するための方法として、音声データの分析からレッスン中に発せられる音声情報の同定とその区間検出を実施した。一般に、演奏指導を実施する際に考慮する必要がある主要素は、演奏音と発話である。これらを統合的に組み合わせることによって、演奏者は独自の演奏表現を構築する。楽器演奏における教授学習の相互行為を記述した研究では、指導者と学習者両者の発話と演奏音に着目した分析が主流となっている [Paul 94, Suchman 87, Sudnow 78]。これらの先行研究に倣い、本研究で対象とする音声データは、演奏音 (music)、発話 (speech)、ノイズ (noise)、無音 (noEnergy) の4つの要素に着目した分析をおこなった。

まず、演奏音に関して、一般的にクラシックギターの

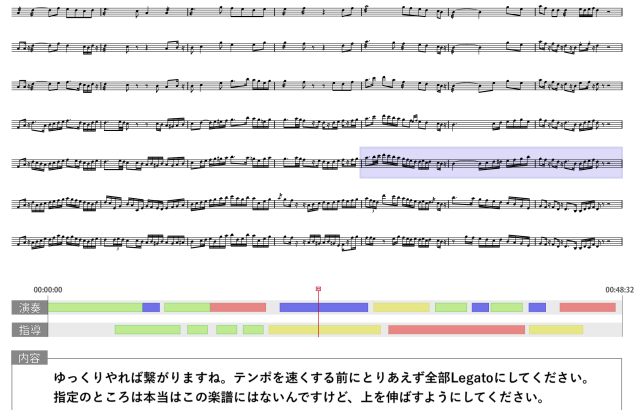


図1 本研究で開発するシステムの想定図

指導現場では、指導者と学習者のためにギターが2台必要となる。2つの楽器音の違いとして、楽器の音響的な特徴や演奏者の弾き方などが複雑に混ざっている。次に発話に関して、個人の特徴があるが、声道長や声帯などの先天的な発声器官の個人差によるものと、アクセントなどの後天的な話し方のくせによるものが相互にからみ合っている。上述以外の音声として、メトロノームや環境音などがノイズの一例である。また、クラシックギターには非常に多くの演奏法があり、中でも効果音楽法^{*1}は、楽器の胴体をタップしたり弦を引っ掻いたりするため、通常のクラシックギター音とは区別されたり雑音として扱われることが予想される。このような多種多様な音声混在するデータから、音声情報と演奏区間の同定によって指導者と学習者の演奏活動を正確に記述することで、より効果的なレッスン指導への応用や、指導者と学習者の演奏構築過程の解明に寄与できる。

3.1 対象データ

本研究で対象とするレッスン形態は、クラシックギターのマンツーマンレッスンである。指導者はプロの演奏家として活動している成人女性であり、国内トップレベルのギターコンクールを複数優勝した経験を持つ。指導歴は7年で大学生から高齢者まで幅広い年齢層の指導にあたっている。学習者はアマチュアの成人男性であり、コンサートやアマチュア向けのギターコンクールの出場に向けて積極的に取り組んでいる。このペアの受講歴は5年と比較的長く、良好な関係が構築されている。

音声データは、学習者が個人の録音機材 TASCAM

^{*1} ギター奏法オントロジー [Iino 19] で定義されている、効果音を出すことを目的とした演奏法。<https://github.com/guitar-san/Guitar-Rendition-Ontology>

表1 対象データと音声区間分析結果の一覧

ID	楽曲 / 作曲者	日付	時間	演奏 (music)	発話 (speech)	ノイズ (noise)	無音 (noEnergy)
01	El Testament d'Amelia / F. Tarrega	2018/07/19	22:02	55.9%	29.5%	12.9%	1.7%
02		2018/07/26	18:42	62.7%	27.0%	2.7%	7.5%
03		2018/08/03	25:20	91.6%	6.1%	1.5%	0.8%
04		2018/08/09	28:56	84.0%	12.0%	1.3%	2.7%
05	Introduction and Variations on a Theme by Mozart / F. Sor	2018/07/26	14:06	77.7%	17.2%	0.5%	4.6%
06		2018/08/03	18:25	94.0%	2.1%	1.1%	2.8%
07		2018/08/09	25:16	79.5%	15.3%	0.9%	4.3%
08	Olite from Castillos de Espana / F. M. Torroba	2018/10/04	15:13	78.0%	17.5%	2.1%	2.3%
09		2019/01/17	24:48	86.3%	5.7%	8.0%	0.0%
10		2019/02/15	28:10	87.5%	10.0%	0.4%	2.1%
11		2019/03/08	21:23	93.9%	2.3%	0.7%	3.1%
12		2019/05/30	31:15	95.8%	1.8%	2.2%	0.1%
平均				82.2%	12.2%	2.9%	2.7%

DR-05 を用いて約 1 年間に及ぶレッスンを記録した結果であり、これまでに約 100 件以上のデータを収録している。本稿では、その中から本著者らがそれらを楽曲ごとに分類し選定したものである。具体的には、1 曲につき 3 回以上のレッスンを実施した次の 3 曲を対象とし、全 12 件分の WAV データの分析を行った；(a) El Testament d'Amelia (アメリカの遺言)、(b) Introduction and Variations on a Theme by Mozart (魔笛の主題による変奏曲)、(c) Olite from Castillos de Espana (スペインの城よりオリーテ)。

3.2 分析手法とその結果

対象とする音声データの解析手法として、Doukhan が提案した inaSpeechSegmenter[Doukhan 18] を用いて、音声情報の同定とその区間検出をおこなった。inaSpeechSegmenter は、男女の発話時間割合推定に基づく大規模な男女平等研究を行うために設計された CNN ベースの音声区間検出フレームワークであり、音声信号を音楽、音声、雑音の均質なゾーンに分割することができる。inaSpeechSegmenter による分析から、先述した演奏音 (music)、発話 (speech)、ノイズ (noise)、無音 (noEnergy) の 4 要素に分類した結果を表 1 に示す。

本分析結果では、music が音声情報の中で最も割合が大きかった要素であった。一方で speech は非常に小さく、一部は noise や noEnergy とほぼ同等であった。本結果を概観すると、各レッスンの初期段階 (ID:01, ID:05, ID:08 など) においては、発話区間 (speech) の割合が多く、レッスン回数を重ねるごとに (ID:03, ID:06, ID:12 など) 減少していくという一定の傾向が見られた。これは、教授学習場面の序盤では、口頭による指導が多いが、終

盤になるにつれて実演する機会が増していくことが推測できる。このように音声区間分析の結果から、特定の楽曲を熟練する過程で発話と演奏のインタラクションやそのパターンに変化があることが伺える。一方で、技術的課題として、本分析では演奏中 (music) の発話 (speech) は全て music としてタグ付けされるため、演奏と同時に発せられた発話について考慮されていない可能性があると考えられる。また、対象データでは、指導者と学習者の歌唱頻度が高かったことも、その結果、music の割合が大きくなった一要因である考えられる。

音声区間の分割において、music の割合が大きかったことを受けて、手動による書き起こし結果をもとに音声区間の再分類を行った。図 2 は、ID: 03 の結果を可視化したものである。図内の最上部は inaSpeechSegmenter によって分割された項目の変化、中央は指導者の変化を、最下部は学習者の変化をそれぞれ示している。青色は music、赤色は speech、水色は noise と noEnergy である。薄い赤色は、inaSpeechSegmenter では music に分類されたが、実際にはその間に発話をしてきた箇所として、music&speech と分類した。また黒色は、inaSpeechSegmenter で speech に分類されたが、発話を聞いていた場合を表しており、no-speech と分類した。

図 2 の結果から、指導者と学習者の変化に着目すると、music&speech の割合が非常に多く見られた。特に多くの音楽レッスンでは発話が大半であり、学習者は傾聴と演奏を繰り返している。しかし今回の結果からは、学習者の music&speech も非常に多く、活発な会話が行われていたことを示している。これは、両者が互いに発言しやすい良好な関係が築けていることを示唆している。このように、音声区間の情報が指導者と学習者の関

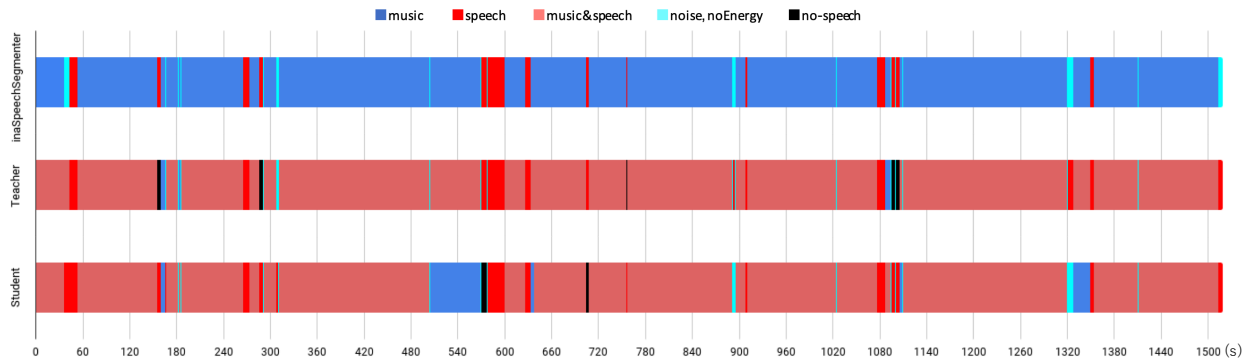


図2 ID:03 (El Testament d'Amelia) を対象とした音声区間分析の結果

係性を表す一つの指標になる得ることが推測される。

4 書き起こしデータに基づく発話分析

教授学習場面における発話と演奏の状況やそのパターンを明らかにするため、3.1節で詳述した音声データ全12件を対象に、発話の書き起こしとその分析をおこなった。書き起こし作業は筆者ら2名が実施した。各発話は、前後の発話との時間間隔や意味的なまとまりに考慮した上でセグメント化し、発話IDや発話者、発話の開始/終了時刻、発話長などの情報も同時に記録した(表2)。

4.1 形態素解析結果から見る教授学習場面の傾向

書き起こしたID:01からID:04までの4件のデータに対して、形態素解析を行った。その結果、名詞は、指導者：523, 168, 217, 268, 学習者：510, 166, 124, 91, 動詞は、指導者：445, 132, 190, 213, 学習者：377, 95, 86, 49, 形容詞は、指導者：110, 43, 72, 66, 学習者：68, 20, 17, 16 (ID: 01~04の順)の単語数が抽出された。顕著な特徴としては、名詞に分類された単語の中で「これ」「それ」といった指示語が多くみられたことである。名詞の約20%が指示詞であることがわかった。これは、楽譜を用いた指導であることや、弾き方や音色に関する内容が多いためだと考えられる。特にID:01では、音色の話から「爪の削り方^{*2}」に関する指導が行われ、爪の箇所や形を表す際に指示詞を多く用いた。このように、

^{*2} クラシックギターは、基本的に演奏者の地爪で弦を弾くため、爪の形状や磨きの細かさによって音色が変化する。

表2 ID:03 (El Testament d'Amelia) を対象とした書き起こしデータと意味ラベルの付与

ID	発話内容	意味ラベル
43	あと、タッタタって切れるのはやめましょう。	Intonation
44	あくまで、レーファーっていう旋律なので、その中にレーファーって息吸うことはないと思います。	Intonation
45	あとこう振りかぶらないで。	Technical Facility
46	弾けなくなっちゃう。	Technical Facility
97	でもあのスピードで弾けるようになってるのであれば本番ね、あの、上がってもそこまでしかいかないので。	Tempo
98	あの、だからこれからはちょっと落として弾くようにして、上がってもそれくらいになるようにしていきましょう。	Tempo
99	あとね、なんかね、ちっちゃくなったりするので、もうちょっとはつきり一個一個出していいと思います、音がちっちゃい。	Dynamics

指示詞を抽出することで、楽譜のどの情報に対して発言されているのか、あるいはどのような指導が行われているのかを推測できる可能性がある。

また、重要単語の同定として、1レッスンデータごとに含まれる名詞と形容詞を抽出し、TF-IDF法の適用によって得られる単語群の上位10%を抽出した。分析結果から、各レッスンの初期段階(ID:01, ID:05, ID:08など)では、“伸び、溜め、雰囲気、ストーリー、コント

ロール”といった単語が得られる一方で、レッスン回数を重ねるごとに (ID:04, ID:07, ID:12 など) “Harmonics, Vibrato, Portamento, Ceja” といった音楽用語やクラシックギター奏法などの固有表現を含む単語が上位に選定されることが分かった。これは、広義の意味を含む内容から、より詳細かつ専門的な技法に関する指導へと教授内容が変化していることが推測できる。このように、形態素解析によって、レッスン固有の特徴や傾向を正しく理解するための手掛かりになると考えられる。

4.2 教授場面における意味ラベルの推移

本研究で取り扱うレッスンでは、具体的な指摘や助言などは、主に指導者の発話に現れる。そのため、本研究では、各発話の内容を把握する指標として、先行研究 [Ellen 03] を参考に、指導者の各発話に対して、テンポ (Tempo)、技法 (Technical Facility)、音楽の表現 (Intonation)、リズム (Rhythm) 音の強弱 (Dynamics)、その他 (Other) の 6 項目の意味ラベル (以下、“意味ラベル”と呼ぶ) を手動で付与した (表 2)。本ラベルの付与により、教授学習場面において、どのようなタイミングでこういった指示が行われているのかを定量的に判断することができる。

ID:01~04 (El Testament d’Amelia) を対象とした意味ラベルの付与結果として、表 3 に意味ラベルの割合と、図 3 に全体の推移を示す。表 3 より、テンポ (Tempo) に関する意味ラベルが 32.6% と最も割合が大きかった要素であった。続いて、音楽の表現 (Intonation) の意味ラベルが 31.5% と高い割合で出現した。一方で、音の強弱 (Dynamics) についての割合は低く、全体を通して教授されることが少ないことが分かった。また、図 3 の結果から、音楽の表現 (Intonation): 10, 5, 65, 49, リズム (Rhythm): 1, 7, 1, 21 (ID: 01~04 の順) とレッスン回数を重ねるごとにその値が増加していくことが分かった。この結果から、全体的な傾向として、技術的な教授内容から、より音楽的な表現や全体のリズム感を練っていく教授内容に変化していることが示唆される。以上より、意味ラベルの付与とその解析によって、複数のレッスンを横断する演奏構築過程の解明に寄与できると考えられる。

5 おわりに

本稿では、クラシックギターの指導時の収録音源を分析し、指導現場における様々な特徴や問題点について議論した。具体的には、畳み込みニューラルネットワーク (CNN) ベースのフレームワークを用いて音声区間の検

表 3 ID:01~04 (El Testament d’Amelia) に含まれる意味ラベルの割合

意味ラベル	回数	割合
テンポ (Tempo)	133	32.6%
技法 (Technical Facility)	77	18.8%
音楽の表現 (Intonation)	129	31.5%
リズム (Rhythm)	30	7.3%
音の強弱 (Dynamics)	13	3.2%
その他 (Other)	27	6.6%
合計	409	100.0%

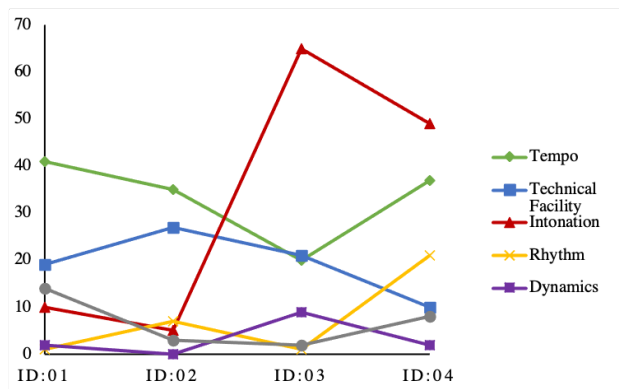


図 3 ID:01~04 (El Testament d’Amelia) を対象とした意味ラベルに関する推移

出とその可視化を行った。さらに実際の音声データから発話を書き起こし、形態素解析による会話分析、そして意味ラベルの付与とその解析により、楽器指導における現場固有の特徴や傾向を提示することができた。

今後、第 2 章で先述したシステムや協働的な演奏構築過程の可視化に向けて、(a) 一人称研究 [諏訪 13] と身体知研究 [古川 06] のような個の事象に焦点を当てたより主観的な研究、(b) 機械学習のような客観的かつ普遍的な研究を、並行して進めていく。これにより、様々な事象に対応しうる知識の体系化を目指す。

参考文献

- [Akoumianakis 08] Akoumianakis, D., Vellis, G., Milolidakis, L., Kotsalis, D. and Alexandraki, C.: Distributed collective practices in collaborative music performance, ACM DIMEA: Digital Interactive Media in Entertainment and Arts, pp. 368–375 (2008).
- [Doukhan 18] Doukhan, D., Carrive, J., Vallet, F., Larcher, A., and Meignier, S.: An Open-Source

- Speaker Gender Detection Framework for Monitoring Gender Equality, IEEE International Conference on Acoustics Speech and Signal Processing, ICASSP, Vol. 7, pp. 103–122 (2018).
- [Ellen 03] Ellen, M.: A Descriptive Analysis of Error Correction in Instrumental Music Rehearsals, Journal of Research in Music Education, Vol. 51, No. 3, pp. 218–230 (2003).
- [古川 06] 古川康一: 身体知としての弦楽器演奏スキル, バイオメカニズム学会誌, Vol. 30, No. 1, pp. 17–20 (2006).
- [Gurevich 06] Gurevich, M.: JamSpace: A networked Real-time Collaborative Music Environment, ACM CHI Extended Abstracts on Human Factors in Computing Systems (2006).
- [Iino 19] Iino, N., Nishimura, S., Nishimura, T., Fukuda, K., Takeda H.: The Guitar Rendition Ontology for Teaching and Learning Support, The 13th IEEE International Conference on Semantic Computing (ICSC), Resource track, Vol. 1, pp. 404–411 (2019).
- [Lim 09] Lim, K. A. and Raphael, C.: InTune: a musician’s intonation visualization system, ACM SIGGRAPH (2009).
- [松原 12] 松原正樹, 諏訪正樹, 斎藤博昭: インタラクティブな楽譜色付けによるオーケストラスコア理解支援システム, 人工知能学会論文誌, Vol. 27, No. 5, pp. 281–295, (2012).
- [Kawase 07] Kawase, S., Nakamura, T., and Draguna, M.: Communication channels performers and listeners use: a survey study, Proceedings of the International Conference on Music Communication, pp. 76–79 (2007).
- [難波 97] 難波精一郎 編集: コンピュータと音声, 音の科学, 朝倉書店, pp. 103–118 (1997).
- [Paul 94] Paul F. Berliner: Thinking in Jazz: The Infinity Art of Improvisation, Chicago, University of Chicago Press (1994).
- [坂本 15] 坂本雄彦: 演奏練習における音声ログの楽譜への対応付けと可視化および演奏に関する議論への活用, 筑波大学大学院博士課程 システム情報工学研究科修士論文 (2015).
- [Suchman 87] Suchman, L. A.: Plans and situated action, New York: Cambridge University Press (1987).
- [Sudnow 78] Sudnow, D.: Ways of the hand, Cambridge, Harvard University Press (1978).
- [諏訪 13] 諏訪正樹, 堀浩一: 特集「一人称研究の勧め」にあたって, 人工知能学会誌, Vol. 28, No. 5, pp. 688 (2013).

決定木と視線特徴量による車酔い自動判定モデルの構築と精度評価

Development of Automatic Recognition Model of Motion Sickness by Decision Tree and Features of Line of Sight and Accuracy Evaluation

奥山祥太^{1*} 大前佑斗¹ 豊谷純¹ 浦田奈愛²

Shohta Okuyama¹, Yuto Omae¹, Jun Toyotani¹ and Nae Urata²

¹ 日本大学大学院生産工学研究科

¹ Graduate School of Industrial Technology, Nihon University

² 三菱自動車工業株式会社

² Mitsubishi Motors Corporation

Abstract: In Japan, where automobiles have become popular enough, many people are suffering from motion sickness. Improvement of our line-of-sight while on the car can be an idea to cope with carsick. However, that possibly depends on individual characteristics. Although it would be helpful if there is a system that presents individualized line-of-sight trends to avoid motion sickness, the research is insufficient. Therefore, the purpose of this study is developing of this system, and we have analyzed the line-of-sight trend by data mining.

1 はじめに

コロナ禍の現在、新しい生活様式[1]が提言され、日々の日常生活から働き方まで徐々に変化してきている。働き方においては、テレワークが主になりつつあり、出社する頻度も少なくなっていることから大都市から地方へと移り住む人たちも少なくない。地方での生活は、都会と比較し長距離の移動が求められるため、多くの世帯が自動車を保有している[2]。そのため、移住をきっかけに自動車を購入する人が増加している。大都市においても、3密回避を目的とした移動手段の1つとして、自動車の需要が高まっている[3]。

このような社会の変化により、以前と比べて自動車に乗る機会が増えた人は多いのではないかと考えられるが、自動車に関する悩みの1つとして、車酔いがあげられる。車酔いは、自動車の誕生から長い月日が経過した現代でも多くの人が悩まされている。そんな車酔いの改善に対する研究は多く存在するが、車酔いの要因とされている前庭器官や体性感覚に着目したものが多く、視覚情報に焦点を当てているものは少ない[4]。そのため、車酔いを改善する視線動向に関しては、前方の車といった動きが比較的穏やかな対象物を中心的に眺めているか、目を閉じていたほうがよい[5][6]と言った大雑把な改善法が多い。視線動向は人によって異なるため、車酔いを解消する眼球の動かし方は一意に定まらないとされていることが、この理由であると考えられる。

そこで、個人に適合した形で車酔いを解消させる視線動向のアドバイスを提示するシステムの開発を行うことが必要である。このためには、精度の高い車酔いの判定モデルを構築することが求められる。さらに、そのモデルを活用し、車酔いする人の視線動向をどのように変化させると酔いにくくなるのかについて明らかにすることが要求される。そこで本研究では、その第一ステップとして、視線から車酔いを自動判定するモデルの構築を行う。

2 乗酔いの原因とメカニズム

「動揺病」とも呼ばれる、乗り物酔いの主な症状としては、悪心、嘔吐、顔面蒼白、冷汗などが挙げられるが[7][8]、現在これらの原因として最も多く取り上げられている仮説は、感覚混乱説 (sensory conflict theory)[9]である。感覚矛盾説とも言われるこの仮説に共通した乗り物酔いの発症モデル[10]を図1に示す。

まず乗り物に乗った際に受ける能動的・受動的な刺激情報が脳に伝達される。主に伝達される情報としては、前庭器官である内耳からの前庭情報、視覚器官からの視覚情報、皮膚や筋肉、関節や内臓器官などの体性感覚からの情報の3つとされている。こうした刺激に対し脳内では、直ちに抗重力筋における前庭-脊髄反射 (vestibule-spinal reflex)による立ち直り反射や外眼筋における前庭眼反射 (vestibulo-ocular reflex)による代償性眼球運動などのポジティブ・フィードバック機能で、身体の平衡感覚を維持

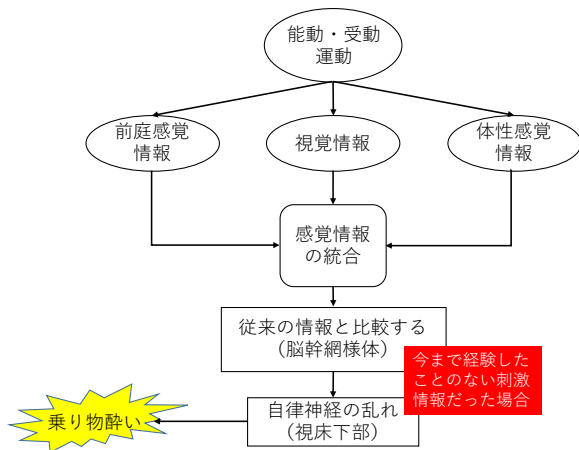


図1 乗り物酔いの発症機序モデル[10]

表1 車酔いに関するアンケート

Q. あなたは車酔いをしますか？	
(1)	必ず酔う
(2)	頻繁に酔う
(3)	酔う時もある
(4)	何度か酔った経験がある
(5)	1度は酔った経験がある
(6)	1度も酔った経験がない

しようとする。それと同時に、その刺激と今まで経験した記憶情報との比較を脳幹網様体で行う。比較をした上で、従来に経験したことのない刺激を感じると脳はその情報を処理できず、不快な状態であると判断する。その結果、脳の視床下部が反応し、血圧や呼吸、内臓の働き全般をつかさどる自律神経の乱れが発生する。これにより、吐き気や眩暈などといった乗り物酔いを引き起こすとされている[10][11]。

3 研究方法

3.1 被験者

被験者は健康な男女37名(男28名,女9名,年齢22±2歳)で、事前に表1のような車酔いに関するアンケートを行った。このアンケートにおいて(1)~(3)に該当する人を車酔いしやすい人、(4)~(6)に該当する人を車酔いしにくい人として分類した。その結果、車酔いしやすい人17人、車酔いしにくい人20人となった。



図2 実験の様子

表2 実験に使用した動画

種類	場所	場面
動画A	幕張本郷	街中
動画B	幕張湾岸	直線道路
動画C	習志野市大久保	住宅街

3.2 実験環境

期間は2020年9月21日~10月30日の40日間で、被験者に検出レート60Hzの視線追跡装置(EMR-9, ナックイメージテクノロジー)を装着させて、着席位置から140~145cm程度離れた位置にある、液晶ディスプレイ(55inch)に映し出された映像を閲覧させ、その間の視線動向を測定した。なお、視点の位置がディスプレイの中心と平行になるように椅子の高さも調節した。また、試験中に頭を動かすことにより視線追跡装置のキャリブレーション精度が下がることを防ぐため、固めの高反発ネックピローを付けてもらい頭を固定した。

使用した動画の内容を表2に示す。街中の映像である動画Aは、緩やかなカーブはあるものの基本的に真っ直ぐな道を走行しているものであるが、対向車、歩行者は共に多く映っているものである。直線道路の映像である動画Bは、道幅の広い3車線の道路の中央車線を走っているものであるため、追い越し車両は複数回見受けられるが歩行者は映っていない。また、植樹された中央分離帯もあるため対向車も映っていない。住宅街の映像である動画Cは、十字路、T字路が多く、頻繁に右折・左折を繰り返しているものである。

これらの映像は、ダッシュボードにカメラ(HERO9 Black, GoPro)を設置し、実際に走行している景色を撮影したものである。こうして撮影した映像から、場面の異なる3パターンの映像(90秒)を作成し各被験者にA→B→Cの順番で2回ずつ閲覧させた。

表3 用いた特徴量とその物理的性質

特徴量名	物理的性質
CX 平均,CY 平均	時系列信号の平均値
CX 分散,CY 分散	時系列信号のばらつき
CX 歪度,CY 歪度	ヒストグラム時の歪み具合
CX 尖度,CY 尖度	ヒストグラム時の尖り具合
累積視点移動距離	視点移動距離の累積
平均視点移動距離	単位時間あたりの視点移動距離

4 分析方法

4.1 視線計測データと事前処理

視線追跡装置を利用して、実験中の視線情報を、時系列信号として取得した。このデータは、ディスプレイの左上を原点として、右に行くごとに X 座標値が増加し、下に行くごとに Y 座標値が増加するものである。その中で、右目と左目の両方の動きから導き出される補正点のデータ (CX, CY) を採用した。実験中の瞬きやよそ見などで生じたエラー値は、全体からエラー値を除いた値の平均値に差し替えた。また、各動画における視聴 1 度目の視線情報を教師データ、2 度目の視線情報をテストデータとし精度評価を行なった。

4.2 判定モデルと特徴量

まず初めに、実験で得られたデータを表 3 に示す、10 次元の特徴量ベクトルに変換した。これを用いて、決定木とニューラルネットワーク (以後、NN) による精度評価を行った。その後、実験環境のわずかな違いが大きく反映されてしまう CX・CY 平均と、フィードバックにおいて原因を考えにくい尖度の 2 つを含めない、6 次元 (CX・CY の分散と歪度、2 種類の移動距離) での分析も行った。最後に、眼球の動かし方に着目した 4 次元 (CX・CY の分散、2 種類の移動距離) での分析を行った。

5 分析結果および考察

表 4 に、決定木を使用した場合の精度を示す。決定木による分析結果では、実験に使用する最適映像は直線道路の場面である動画 B であることが分かる。また使用する特徴量においては、4 次元と 6 次元において同じ最大精度が得られたことから、平均と尖度が判定に寄与しない特徴量であることが分かる。

表 5 に、NN を使用した場合の精度を示す。NN の分析結果では、動画 B を用いて得られた視線データを、特徴量を 6 つ採用した 6 次元で分析した場合に最大精度が得られた。最大精度が得られた時の条件を表

表4 決定木における特徴量・動画別精度

特徴量	動画 A	動画 B	動画 C
10 次元	0.514	0.595	0.676
6 次元	0.649	0.757	0.541
4 次元	0.595	0.757	0.568

表5 NN における特徴量・動画別精度

特徴量	動画 A	動画 B	動画 C
10 次元	0.676	0.622	0.595
6 次元	0.622	0.730	0.622
4 次元	0.649	0.595	0.649

表6 NN における最大精度時(0.730)の条件

中間層	学習係数	学習回数
4	0.001	10000
5	0.001	10000
9	0.01	5000

6 に示す。また、特徴量が 4 次元の場合と歪度を含めた 6 次元の場合では、約 15% も正答率が違うことから、視線情報から車酔いを自動判定するのにおいて、歪度は必要な特徴量であると考えられる。その他に、決定木、NN 共に動画 B での精度が高くなっている理由としては、直線道路という対象物が少ないシンプルな環境のため、比較的穏やかな視線動向であることが挙げられる。視線動向は人によって必ず異なる為、酔う・酔わないに関係なく全体的に誤差が生じる。しかし、直線道路のようなシンプルな映像では、注目する対象物が少ないため全体的に視線動向が似ていて誤差が小さくなる。そのため、その中で僅かな違いなども結果に反映されやすいという観点から、より高い精度が出やすいのではないかと考えられる。

この 2 つの結果より、直線道路の動画 B を用いた視線計測によって取得した X, Y の時系列データを、CX/CY の分散と歪度・累積と単位時間あたりの移動距離の 6 次元の特徴量をもとに、作成した決定木が最適モデルであると考えられる。

次に、最適モデルの構築条件をもとに図 3 の決定木を可視化した。まず 1 つ目の判定基準としては、視点の上下方向の座標の散らばり具合を表す分散 (CY 分散) を指標となる。この判定では、CY 分散が 92.75 以下の人の 11 人中 9 人が酔いにくい人であるという結果になった。この結果から、車酔いしにくい人は、縦方向の視線動向を殆ど行っていないか、行っていたとしても移動幅が小さい傾向であることが分かる。

2 つ目の判定基準としては、視点の左右方向の座

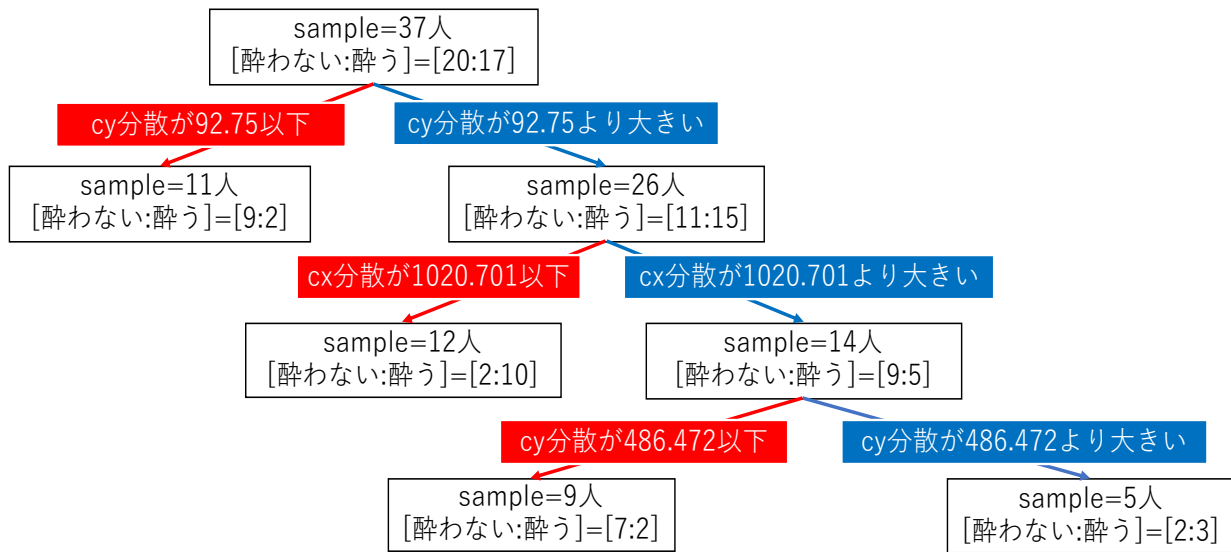


図3 決定木による車酔い自動判定モデル

標の散らばり具合を表す分散 (CX 分散) が指標となる。この判定では、CX 分散が 1020.701 以下の人の 12 人中 10 人が車酔いしやすい人であるという結果になった。この結果からは、横方向の分散が小さい方が酔いやすいと考えられる。しかし通常、極力眼球を動かさないか、目を閉じておいた方が酔いにくい[10]とされている事から、これと相反する結果である。この理由については今回の分析では明らかにできないため、今後詳細に調査することが必要である。

3 つ目の判断基準としては、再び視点の上下方向の座標の散らばり具合を表す分散 (CY 分散) を指標としている。この判定では、分散値が 486.472 以下の人の 9 人中 7 人が車酔いしにくい人であった。この結果からは、最初の判別と同様に、縦方向の視線動向が関係してきていることが分かる。

今回の判定モデルから考えられる、車酔いに関する特徴としては、縦方向の視線動向が比較的緩やかである人は車酔いしにくいということと、横方向の視線動向は車酔いに比較的関係がないのではないかとこの 2 点が挙げられる。今回使用した直線道路の映像は、対向車も映っておらず、追い越す車両も比較的少なかった為、視線が集まる対象としては前方を走る車であった。試験中の視線追跡映像から、前方の車を注視し目で追っている人の視線動向は、車間が縮まるにつれて対象物が大きくなる為、無意識に視線が少し上に移動しているように見受けられる。この視線動向は、車間によって視線の高さが変わることを裏付けている。これを基に推測すると、車酔いしにくい人は、ある一点の対象物を注視

するのではなく、ざっくりとした範囲を見ているのではないかと考えられる。一方で車酔いしやすい人は、ある 1 点を集中して目で追う習性があるのではないかと考えられる。

これらの特徴から導き出せる、車酔いを改善する方法としては、縦方向の視線動向を避けるとともに、ある 1 点にピントを合わせて目で追うことを避けることが挙げられる。

6 おわりに

本研究では、視線情報から車酔いを自動判定するモデルを構築した。また、このモデルから車酔いしやすい人の特徴を導き出し、この特徴に基づいた車酔いを改善させる視線動向を提案した。しかし、視線動向は環境要因によって変化するものであると考えられるため、今回の分析で得られた判定モデルは、使用した映像にのみ適用されるものである。また現状では、機械学習に用いるデータ数が少ないため、より個人に適合した酔わない視線動向を導き出すことはできていない。そのため、より多くの視線データを取得し再検証する必要がある。今回は決定木の分析結果を基に車酔いに対する改善策を考えたが、決定木は可読性が高い一方で、分岐を決定する時に一つの変数しか考慮されないため、判定に重要な特徴量が反映されていない可能性がある。そのため、特徴量が全て反映される NN のような手法を用いて、必要な特徴量を再考することも今後の課題である。

参考文献

- [1] 新型コロナウイルスを想定した「新しい生活様式」の実践例,厚生労働省:

- https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431_newlifestyle.html (最終閲覧日:2020/11/6)
- [2] 自動車保有台数,自動車検査登録情報協会:
<https://uub.jp/pdr/t/cr.html> (令和2年6月末現在)
- [3] コロナ下で中古車販売好調 3密避け需要増、新車納期延び待ちきれず,産経デジタル SankeiBiz:
<https://www.sankeibiz.jp/business/news/201015/bsc2010150500007-n1.htm> (最終閲覧日:2020/11/6)
- [4] 中川千鶴,大須賀美恵子:VE酔い研究および関連分野における研究の現状,日本バーチャルリアリティ学会論文誌,3(2),pp.31-39.(1997)
- [5] N. A. Webb & M. J. Griffin: Optokinetic stimuli: Motion Sickness, Visual Acuity, and Eye Movements, *Aviat. Space Environ. Med.*,73(4),pp.351-358.(2002)
- [6] M. J. Griffin & M. M. Newman: Visual Field Effects on Motion Sickness in Cars, *Aviat. Space Environ. Med.*,75(9),pp.739-748.(2004)
- [7] D. B. Tayler & P. Bard: Motion Sickness. *Physiol. Rev* 29, pp.311-369.(1949)
- [8] K. E. Money: Motion Sickness. *Physiol. Rev*50, pp.1-39,(1970)
- [9] J. T. Reason & J. J. Brand: Motion Sickness, Academic Press, London. (1975)
- [10] 平柳要: 乗り物酔い(動揺病)研究の現状と今後の展望. *人間工学*,42(3),pp.200-211.(2006)
- [11] 松永亨,武田憲昭: 動揺病と宇宙酔い. *耳鼻臨床* 81, pp.1095-1120.(1988)

決定木を活用した消費行動を規定する要因の分析について

Analysis of Factors Regulating Consumption Behavior Using Decision Trees

林田 優¹ 豊谷 純¹ 大前 佑斗¹

Yu Hayashida¹, Jun Toyotani¹, Yuto Omae¹

¹ 日本大学生産工学部 マネジメント工学科

¹ Department of Industrial Engineering and Management, College of Industrial Technology, Nihon University

Abstract: It has been proposed for how advertising is used for psychological and attitude changes at various promotional marketing models. Each model emphasizes the importance of spontaneous search and contact with information. This paper shows the differences in consumer values among social networking sites and examines the factors that determine the effects of social networking sites and consumer values on consumer behavior.

1. はじめに

近年、主にネット広告の効果を調べることができ技術が確立してきた影響もあり、広告の効果測定を意識したマーケティング戦略が立てられるようになった。広告の効果測定としては、コミュニケーション効果と売り上げ効果の二つが挙げられる。

コミュニケーション効果の測定の中で、近年では、AIDMA (アイドマ) という生活者の意思決定プロセスが取り上げられることや電通の提唱した AISAS (アイサス) [1] , 日本プロモーション・マーケティング協会の RsEsPs (レップス) [2] など様々なプロモーション・マーケティングが提唱されている。これらは広告が何人にリーチし、表示回数を何回にすればどれだけの認知を獲得するかといった心理や態度変容に及ぼす影響を明らかにするものである。どのモデルの中でも、「情報」との接触や拡散、発信などが重要視されている。

情報の発信や接触に対する先行研究として、インターネットの発達によって、従来指摘されていた刺激-反応型と情報処理型に加え、情報発信を考慮した意思決定プロセスの存在が示唆されている[3]。ファッションの購買に対する先行研究として、若者の社会生活要因の一つである SNS の利用時間の長さがファッションの知識・意識・行動の全てに影響を与えることが示されている[4]。他にも具体的な SNS と消費価値観、スマホ決済に焦点を当てた研究では、ファストファッションの購買を規定する要因が SNS 別によって違いがあることなどが示されている[5]。

流行や興味の移り変わりが激しいファッションにおいて、どのように「情報」との接触や拡散発信をしているのかを若者を中心として分析しているものは多いが、若者以外を対象として行っている研究は

少ない。また米澤[6]は、時代の移り変わりによって、ファッションに対する購買価値観が変化していること、時代はファッションからライフスタイルの時代になったことを指摘している。そのため本研究では、ファストファッションの購買において、利用している SNS、消費価値観、年齢、性別の関係について示唆を得ることを目的に、分析を実施した。

2. 評価項目

人の購買行動についての先行研究として、企業が発信する内容よりも消費者が発信する情報の方が商品の購買に対して良い影響を与えることを示唆している研究[7]がある。このことから、ファストファッションの購買を規定する要因を探るため、SNS や動画投稿サービスなどの WEB 媒体の利用と消費価値観に焦点を当てて検証を行う。

また、若者とそれ以外の購買を決定する要因の違いを分析するために、分析対象を 20 代と 50 代に分割し分析を行う。ここで、年齢によってファストファッションの購買を規定する要因の差異を WEB 媒体の利用の有無と消費価値観から説明することができれば、これらに介入する施策を実施し、より有効的な情報発信や情報との接触を生むことができる。

3. 分析方法

3.1 概要

本研究では、野村総合研究所インサイトシグナルマーケティング分析コンテスト[8]によるデータを用いて分析を実施した。これはマーケティングに関わる多様な質問項目で構成されたアンケート結果である。本研究では使用するデータを以下のように定

めた。まず、被験者の基本的な属性として年齢と性別を使用した。年齢については 20~29 歳, 50~59 歳, 性別については, 男性と女性と定める選択肢を参照した。

WEB 媒体の利用については, w1:Twitter, w2:Facebook, w3:Instagram, w4:LINE, w5:YouTube, w6:Amazon, w7:楽天, w8:Google, 上記 8 項目の利用頻度についての回答結果を使用した(選択肢: ほぼ毎日/週 4, 5 回/週 2, 3 回/週 1 回程度/月 1 回程度/月 1 回未満/利用していない)。この中で, 週 1 回以上利用しているもの(ほぼ毎日/週 4, 5 回/週 2, 3 回/週 1 回程度)を WEB 媒体の利用あり, 月 1 回程度以下(月 1 回程度/月 1 回未満/利用していない)の利用を WEB 媒体の利用なしとラベル付けした。

消費価値観については, s1:とにかく安く経済的なものを買う(以降, 経済的), s2:多少値段が高くても, 品質のよいものを買う(以降, 品質重視), s3:使っている人の評判が気になる, s4:多少値段が高くても, アフターサービスが充実している方がよい(以降, AF), s5:できるだけ長く使えるものを買う(以降, 耐久性), s6:環境保護に配慮して商品を買う(以降, 環境保護), s7:安全性に配慮して商品を買う(以降, 安全性), s8:使い捨て商品をよく買う(以降, 使い捨て商品), s9:プライベートブランド(小売店が独自に販売しているブランド)をよく買う(以降, PB), s10:商品を買う前にいろいろ情報を集めてから買う, s11:有名な人がよいと言っているものを選ぶことが多い(以降, 有名人追従)など, 上記の 11 項目の回答結果がある(選択肢:ない, ある)。

ファストファッションの利用について問う項目として, 週に 1 回以上, 月に 1~2 回程度, 2 ヶ月に 1 回程度, 半年に 1~2 回程度, 年に 1 回程度, ほとんど利用していない, 上記 6 項目がある。

今回行った分析でのターゲットであるファストファッションのラベリングは, 「週に 1 回以上, 月に 1~2 回程度, 2 ヶ月に 1 回程度」と回答した被験者は, 季節を通して定期的に購入をしている可能性が高いと判断し, ファストファッションの利用ありとした。そして, 「半年に 1~2 回程度, 年に 1 回程度, ほとんど利用していない」と回答した被験者は, 頻度が少なく不定期な購入の可能性が高いため, ファストファッションの利用なしとラベリングした。

3.2 WEB 媒体によるファストファッションの購買を規定する要因

被験者の総数は 2500 名分であったが, 前述したアンケート項目の WEB 媒体の利用の有無と年齢性別についての質問項目において, 1 項目以上の回答に欠

損がある被験者を除外した。その結果, 2428 名分が利用可能なデータであった。今回は, このデータの 20 代と 50 代の層別にデータを分けて, 分析を行った。これらの取得したデータに対して, 性別の WEB 媒体とファストファッションの購買を規定する要因を探るべく, 決定木分析を行った(ジニ係数最小化基準, 最大深度 4)。その結果を図 1 から図 4 に示す。

図 1 より 20 代男性では, Instagram の利用の有無によって, ファストファッションの利用率が大きく変化していることがわかる。ルートノードを見ると 20 代男性全体(買わない 121 名, 買う 74 名)の約 38% がファストファッションを買っている。その下段について, Instagram を利用している人は 53.5% (ルートノードに比べ 15.6% 上昇)の人が購入し, 利用していない人は 29% (9% 減少)の人しかファストファッションを購入していない。また Instagram・Google・Amazon を利用しているユーザー群は, 利用率が全体(ルートノード)の 38% から図 1 の右下のように約 70% (37 名中 26 名)にまで向上することがわかる。そのため, Google の検索エンジンを利用した検索連動型広告や Instagram の広告を利用することにより, より効果的な広告効果を得られることができると考えることができる。次に, 図 1 の最下段右から 2 つ目, Instagram と LINE は利用するが, Google を利用しない人は, ファストファッションを利用しないことから, プライベートブランドで展開しているファストファッション以外のファッションを利用している可能性がある。

図 2 より 20 代女性では, ファストファッションを利用している人は, 全体(ルートノード)で 51.4% (212 名中 109 名)だったのに対して, Twitter・Facebook・楽天の利用者は, 図 2 の右下のように 72% に向上する。逆に Twitter と Instagram の利用をしていない人は, 図 2 の上から 3 段目の左端のように約 22% まで利用率が落ちている。このことから, 20 代女性をターゲットとした場合は, Twitter や Facebook に広告を出すとともに, 楽天市場などに出店することが有効的と考えることができる。

一方で, 50 代男性の図 3 は, 20 代では SNS の利用が一番上のリーフノードに現れていたのに対して, Amazon の利用の有無が一番上のリーフノードとして, 現れているのが特徴的といえる。50 代男性の場合, 元々ファストファッションを購入する層は, 全体 283 名中 69 名の 24% しかおらず, 総じてファストファッションの利用率が低いことが窺える。

50 代女性の被験者は, 全体(ルートノード)では約 37% の被験者がファストファッションの利用をしていたのに対して, Instagram を利用している人は,

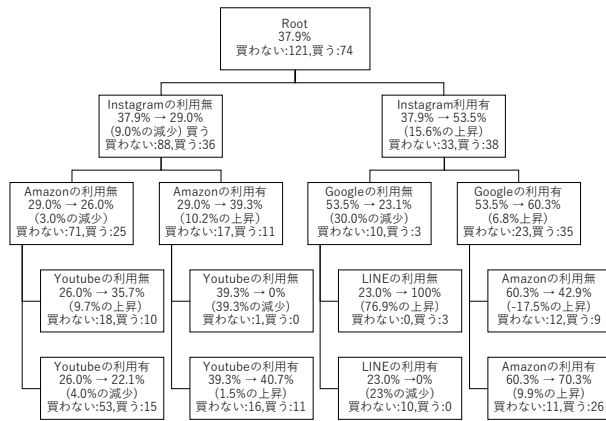


図1 20代男性 WEB媒体のみ

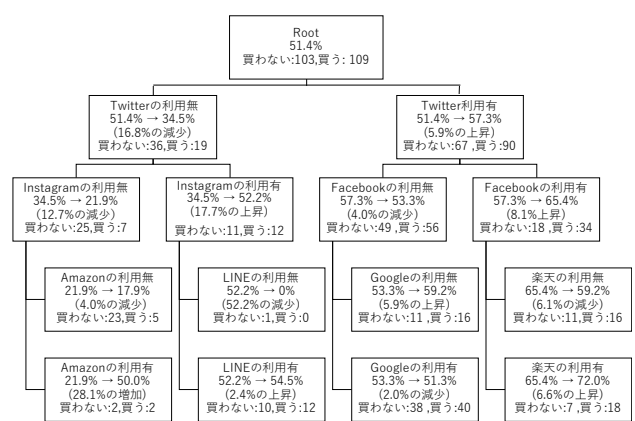


図2 20代女性 WEB媒体のみ

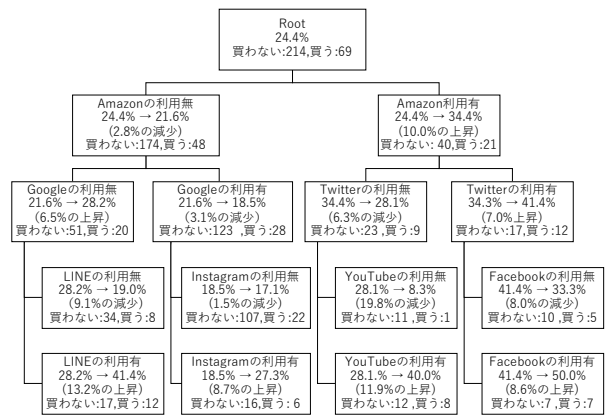


図3 50代男性 WEB媒体のみ

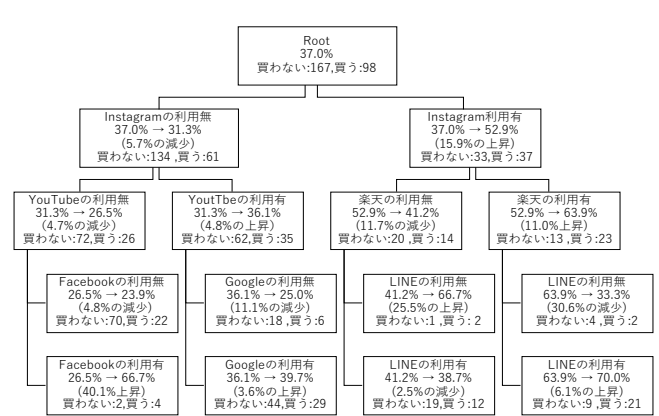


図4 50代女性 WEB媒体のみ

図4の上から2段目の右のようにファストファッションを購入する人が52.9%と約16%上昇している。特にInstagram, 楽天市場, LINEを利用している場合は, 70%がファストファッションを利用していることから, 50代女性をターゲットとした場合, InstagramやLINEに広告を出すとともに, 楽天市場への出店が有効的であると考えられる。一方で, Instagram, YouTube, FacebookなどのWEB媒体の利用がない場合, ファストファッションの購入率が約13%落ちることから, 情報との接触をあまりしない層は, ファストファッションに興味が無いことが窺える。

3.3 WEB媒体と消費価値観によるファストファッションの購買を規定する要因

次にSNSなどWEB媒体にさらに消費価値観を加えた際の調査を行った。3.2節で利用したアンケート項目に加え, 消費価値観についての項目を追加した。これらの質問事項に1項目以上の回答に欠損がある

被験者を除外した。その結果, 利用可能なデータ数は2417名分であった。3.2節と同様に, これらのデータを20代と50代に分割し, 決定木分析を行った(ジニ係数最小化基準, 最大深度4)。その結果を図5から図8に示す。

図5に示す20代男性WEB+消費価値観によると, 図1と同様にInstagramの利用の有無が一番上の条件になっている。そのため, 20代男性の購買行動の意思決定には, 消費価値観よりもInstagramが重要であることがわかる。また図5の20代男性の意思決定として, Instagramを利用しない, プライベートブランドも利用しないがLINEは利用する人(買わない2名, 買う7名)とInstagramやGoogleは利用して, 品質を重視する人(買わない3名, 買う13名)でのファストファッションの利用率が共に約80%と高くなっている。意思決定に安価であることが売りのファストファッションにおいても, 品質を重視する被験者が多いことをから, 20代男性に対して, ただ安いだけでは訴求ポイントになりにくいという問題が示された。

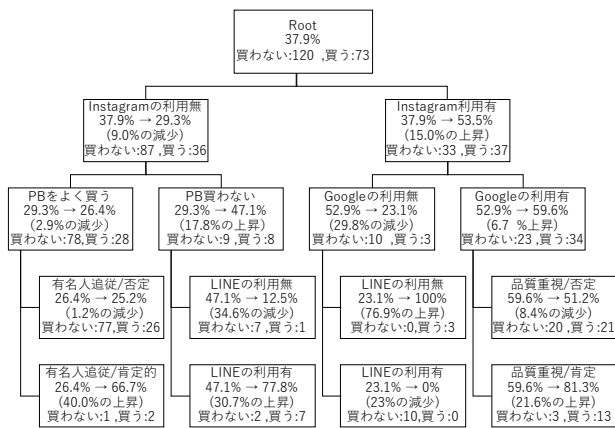


図5 20代男性 WEB+消費価値観

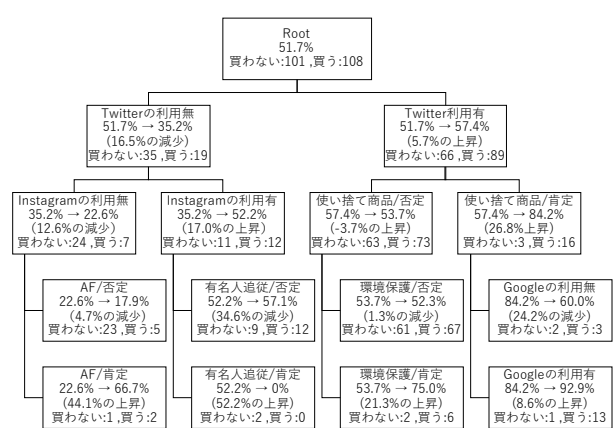


図6 20代女性 WEB+消費価値観

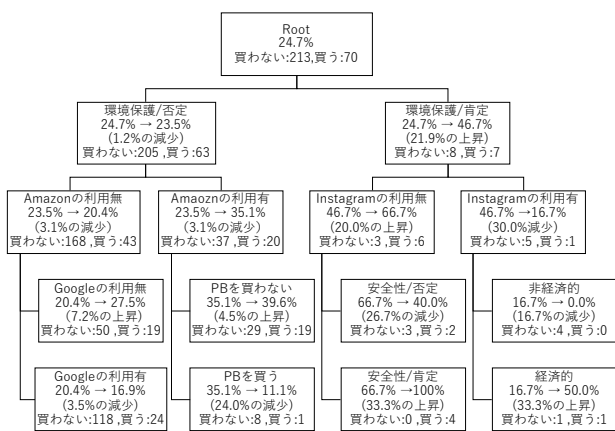


図7 50代男性 WEB+消費価値観

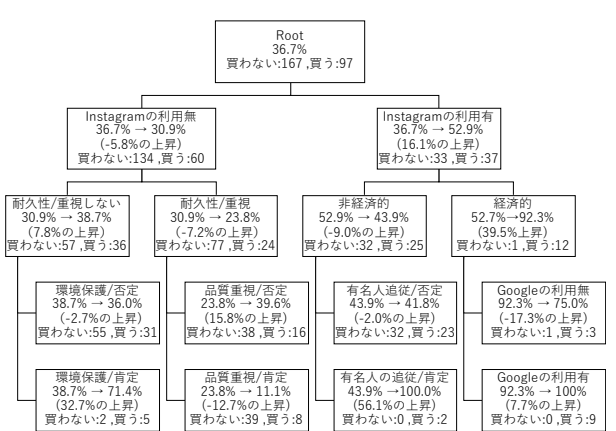


図8 50代女性 WEB+消費価値観

図6の20代女性WEB+消費価値観を見ると、図2の時と同様にTwitterの利用の有無が一番上の条件になっている。そのため、20代女性の購買行動の意思決定には、消費価値観よりもTwitterが重要であることがわかる。また、Twitterの次のノードでは、使い捨て商品であるかどうか重要であると示されている。使い捨て商品であるかどうかは、「ワンシーズン楽しめれば良いと思う」「飽きたらすぐに捨てても惜しくない」などの使い捨てにつながる因子であることが明らかにされている[9]。また、「次々変わるトレンドに使い捨て感覚で消費するもの」や「ブランドに関係なく旬のものの組み合わせを楽しむもの」などの視点を持った消費行動にシフトしつつあるとする先行研究[10]もある。このことから20代女性には、TwitterやGoogleに広告を出してファストファッションの中でも使い捨て商品という点を訴求すると良いと考えられる。また、図2と同様にTwitterとInstagramを利用しない人(買わない24名、買う7名)はファストファッションを購入しない傾向があることがわかる。このことから、Twitterと

Instagramを利用しない被験者は、消費価値観によって、利用するかしないかはあまり変化しないことがわかる。このことから、20代女性におけるファストファッションを利用するかしないかの大きな要因として、SNSであることがわかる。

図7の50代男性WEB+消費価値観より、図3の時と同様に全体的にファストファッションを利用する人が少ないことがわかる(買わない213名、買う70名)。しかし、50代男性にはファストファッションの訴求点として、「環境に配慮して商品を買う」という点をあげることができる。元々の利用率が約24%(買わない213名+買う70名)だったのに対して、「環境に配慮して商品を買う」被験者(買わない8名、買う7名)では、購入率が約47%まで上昇していることがわかる。

図8の50代女性WEB+消費価値観より、図4と同様にInstagramの利用の有無が一番上の条件になっている。そのため、50代女性の購買行動の意思決定には、消費価値観よりもInstagramが重要であることがわかる。また、消費価値観を加えたことによ

り特徴的なものは、50代女性のファストファッションの購買を規定する要因として、経済性が重要だとわかる。経済的であることで、利用率が当初53%(70名中37名)から92%(13名中12名)まで上昇している。このことから、50代女性をターゲットとしたファストファッションでは、Instagramで、安価であることを訴求することが重要であると考えられる。

4. おわりに

今回の分析から、ファストファッションの場合、20代と50代女性に対してInstagramやTwitterなどのSNSの利用が重要であることがわかった。そのため、プロモーション・マーケティングの中で提唱されている「情報」との接触や発信が重要であることが伺うことができる。

しかし、50代男性に対してはWEBの利用が重要でなかったことから、必ずしも全年齢を通じてWEBを通じた情報との接触や発信が重要であるとは限らないことが確認された。そのため、インターネット広告の費用がテレビ広告の費用を上回った現在でも、広告を打ちたいターゲットによっては、WEB経由では情報との接触を得ることが難しいという問題点が示された。今後はさらに深く調査を行なうことで、広告をより効果的に訴求できる戦略を立案するための分析を実施していく予定である。

謝辞

データの利用および学会発表を許可した野村総合研究所インサイトシグナルに感謝する。

参考文献

- [1] 近藤史人, AISAS マーケティング・プロセスのモデル化, JSD 学会誌システムダイナミックス, vol. 8, pp. 95-102, 2009.
- [2] 日本プロモーション・マーケティング協会(編), プロモーション・マーケティング ベーシック, 宣伝会議, 2019
- [3] 清水聡, クロスメディアコミュニケーションと消費者行動, AD STUDIES Vo126, 2 16 -18, 2008
- [4] 渡辺裕子, 町田欣弥, 大学生のファッションの購買意識・行動とソーシャルメディアの影響, 駿河台経済論集, 2020
- [5] 藤原美佳, ファストファッションの消費行動を規定する要因, 繊維製品消費科学, 61 巻 4 号, 2020
- [6] 米澤 泉, 「くらし」の時代: ファッションからライフスタイルへ, 勤草書房, 2018

- [7] 清水麻衣, CGM が消費者の購買意思決定プロセスに及ぼす影響, 商学論集, 第81巻第3号, 2013
- [8] 野村総合研究所(インサイトシグナル), マーケティング分析コンテスト 2020, <https://www.is.nri.co.jp/>.
- [9] 大枝 近子, 佐藤 悦子, 高岡 朋子若者のファストファッション意識に関する調査, 日本家政学会誌, 64 巻 (2013) 10 号, 2013
- [10] 藤原 美佳, 福森 護, ファストファッションにおける消費行動の分析, 繊維製品消費科学, 59 巻 12 号, 2018