

ウェアラブルセンサによる歩行データからの転倒動作検出の試み Detecting Fall Motion from Gait Data Using a Wearable Sensor

大井 健矢^a 松下 光範^{a,*} 堀 寛史^b 高岡 良行^c
Kenya Oi Mitsunori Matsushita Hirofumi Horii Yoshiyuki Takaoka

a. 関西大学総合情報学部

Faculty of Informatics, Kansai University

b. びわこリハビリテーション専門職大学理学療法学科

Faculty of Physiotherapy, Biwako Professional University of Rehabilitation

c. 株式会社 PTS

PTS Co., Ltd.

Abstract: This study aims to develop a system that can detect a person's falling on the ground using a wearable sensor. As the first step of the study, we propose a method to predict the motion types and their directions from time-series signals of acceleration and angular velocity. The proposed method uses the k-shape method to create a classifier to identify the motion details. The classifier discriminated the motions of walking, falling, and stationary from the sensor data. To prevent false detection of falls, we also identified non-fall similar to the fall. The first step in the identification process is to discriminate whether a motion is a fall or a non-fall. If it is a fall, determine the fallen direction at that time. As a result, we revealed that using acceleration data for discriminating fall and non-fall and angular velocity data for discriminating the fallen directions improve the discrimination accuracies.

1 はじめに

我が国の高齢化の進行は深刻な状況にある。2020年10月時点の65歳以上人口は3,619万人となり、総人口に占める割合（高齢化率）も28.8%となった*1。それに伴い一人で暮らす高齢者（独居高齢者）の数も年々増加しており、2015年時点で592万人（65歳以上人口の17.5%）だった独居高齢者は、2040年には896万人（65歳以上人口の22.8%）にまでなると推計されている*2。

独居高齢者の増加に伴って問題になるのが、家庭内での転倒や転落である。東京消防庁のデータ*3によると、高齢者の事故のうち80.8%が転倒や転落に起因している。高齢者の転倒は骨折などの大きな怪我に繋がりやすく介護が必要になる原因にもなることから、転倒後の早

急な対応が重要となる。特に独居高齢者による転倒は長時間発見されない可能性があり、転倒を迅速に発見し、他者に通報する仕組みが必要である。

こうした背景の下、本研究では独居高齢者の転倒動作検出を試みる。転倒検出方法としては、室内の壁や天井にセンサを設置する方法と、ウェアラブルセンサを用いる方法が考えられる。前者の方法には、ドップラーセンサを用いた手法[7]や放射温度センサを用いた手法[5]などが挙げられる。しかしこれらの方法は、室内にセンサを設置する必要があり運用面に問題がある。そこで本研究では後者の方法を採用することとし、ウェアラブルセンサを用いて転倒発生の検知とその際の転倒方向の識別を試みる。転倒方向を識別する理由は、転倒の原因推定を行う際の手がかりとして将来的に活用できると考えたためである。

* 連絡先：関西大学総合情報学部
〒569-1095 大阪府高槻市霊仙寺町 2-1-1
E-mail: mat@res.kutc.kansai-u.ac.jp

*1 https://www8.cao.go.jp/kourei/whitepaper/w-2021/zenbun/pdf/1s1s_01.pdf (2022年2月28日確認)。

*2 https://www8.cao.go.jp/kourei/whitepaper/w-2021/html/zenbun/s1_2_2.html (2022年1月31日確認)

*3 https://www.caa.go.jp/policies/policy/consumer_safety/caution/caution_009/pdf/caution_009_180912_0002.pdf (2022年1月31日確認)

2 関連研究

この章では、センサを活用して人の状態を推定する試みについて述べる。

高橋らはセンサを杖に装着し、歩行動作の検出を試みた [6]。歩行リハビリテーションを行っている高齢者の歩行能力を計測、評価するため、杖に取り付けたセンサの値から対象の歩行状態を推定し、歩行検出については 95.6%、歩行距離については 88.06% の精度で各々推定可能であることを示した。

Vahid らは加速度計を用いて健康に関する日常的な動作の識別を試みた [2]。加速度計を装着した被験者にウォーキング、サイクリングなど健康的な活動をしてもらい、機械学習を用いて分類した結果、88% の精度で識別可能であることを示した。

磯らはアンビエントセンサを用いてテレワーク作業者の状態推定を行った [4]。この研究では、距離センサを部屋に複数設置し、そこから得られるデータとマイクから得られるデータを k-means 法を用いて分類し、会議中とその他の作業を分類可能であることを示した。

Jennifer らはスマートフォンの加速度計を用いて日常的な行動の認識を行った [3]。29 人のユーザに対してウォーキング、ジョギング、登山などの日常的な活動を行わせ、歩く、ジョギングする、階段を昇降する、座る、立つについての予測を行った。その結果、階段を昇降する動作のみ 77.6% だったが、ほかの動作についてはすべて 90% を超える精度で認識できることを示した。

Debraj らはウェアラブルセンサを用いて複雑な動作の区別を試みた [1]。複雑な動作として、床に座る、ソファに座る、ソファに横になる、トイレに座るなどの 19 動作を対象にデータを収集し、それらを機械学習を用いて分類した。その結果、自宅での 19 動作について 80% の精度で分類できることを示した。

3 転倒動作の検出指針

本研究では、ウェアラブルセンサから取得した時系列信号を逐次モニタリングし、装着者の転倒を検知することを試みる。転倒とは、意図せずに地面や床などの低い位置へ倒れたり、ひっくり返ったりすることである。転倒には、つまずきによる転倒、滑ったことによる転倒、バランスを崩したことによる転倒、意識を失ったことによる転倒など、さまざまな原因が考えられるが、本稿ではつまずきによる転倒を対象とする。

ウェアラブルセンサを用いて転倒検出を行う際、転倒



図 1: Vitalgram の外観 (ベルト装着時)

事故発生時にのみ転倒と判定し、歩行などの日常的な動作には誤判定しないようにしなければならない。このとき、意図的に臥位になる動作 (e.g., ベッドで横になる、家具の下を覗き込む) のように、転倒に似た動作 (以下、非転倒動作と記す) も存在するため、誤判定の危険が存在する。そこで、転倒ならびに転倒との誤判定が想定される動作を異常動作、それ以外の日常動作を定常動作とし、段階的に推定する手法を提案する。

本稿では、「歩いている途中で、転倒あるいは転倒に似た動作を行って、そのまま静止した」という行動を想定する。この一連の行動に含まれる動作のうち、定常動作は歩行動作と静止状態の 2 種類からなるものとし、異常動作は歩行から静止に至る際のきっかけとなる動作として、つまずいて転倒する動作 (転倒動作) と、膝をついて意図的に臥位になる動作 (非転倒動作) の 2 種類とし、センサ情報からそれらの識別を行う。

3.1 使用するセンサ

本稿ではアフォードセンス社の Vitalgram を使用する。Vitalgram は日常生活を送りながら生体情報や環境情報を計測できる小型のワイヤレス生体センサである*1 (図 1)。専用のベルトにセンサを固定し、身体に巻き付けて装着することで身体データを取得する。取得できるデータ項目は心電、体表温度、3 軸加速度&角速度、高度、環境温度、相対湿度、熱流の 8 項目であり、これらをもとに、心拍数、心拍間隔、心拍揺らぎ解析による自律神経のバランス指標、呼吸数、熱流計測による深部温度などの項目を推定できる。今回はこのセンサのうち、3 軸加速度&角速度を用いて行う。

安定してセンサ値を取得できる装着方法について予備調査を行った結果、胸部の中央の位置にセンサを装着し、センサおよびベルトをテープで固定する方式が最も安定した結果が得られたため、本稿の実験ではこの装着方法を採用した。

*1 <http://www.affordsens.com/Product-Vitalgram> (2022 年 2 月 28 日確認)

3.2 識別精度の検証法

本稿では分類器による動作の識別精度は次の手順で検証する。

対象者が装着したウェアラブルセンサから時系列信号を取得する。このとき、取得する動作は歩行動作、転倒動作、非転倒動作、静止状態とし、歩行以外の各動作については前後左右の4方向について取得する。得られた時系列信号を機械学習で学習させ、動作の識別を行う分類器を作成し、新たに用意したテストデータを用いて分類器による動作、方向の識別を行う。ここで、テストデータは歩行中に転倒をしてそのまま静止したデータと、歩行中に非転倒をしてそのまま静止したデータの2種類を用いる。識別結果から歩行範囲、異常範囲、静止範囲を設定し、範囲の設定に対する正解率を算出する。このとき、異常範囲における転倒/非転倒動作の識別に対する正解率と、異常範囲、静止範囲における動作方向の識別に対する正解率も算出する。以上3つの識別に対する正解率からどの程度の精度で動作、方向の識別ができたかを考察する。

4 データセットの作成

4.1 データの収集手順

歩行時に異常動作（転倒動作ないし非転倒動作）を行いそのまま静止した、という状況を想定し、データを収集した。したがって収集したデータは、歩行動作、転倒動作、非転倒動作、静止状態の4種類から構成されるデータである。大学生1名を対象とし、センサを付けたベルトを身体に装着することでデータを収集した。センサの装着位置は鳩尾上部の高さの胸部中央位置とし、装着方法はセンサおよびベルトを医療用テープで固定した。歩行動作はセンサを装着した状態で平坦な道を歩くことで取得した。異常動作は、実験時の怪我を予防するために、床に布団を敷いて行った。転倒動作については2, 3歩程歩いた際に、つまずいて転倒した。非転倒動作については同様に2, 3歩程歩き、床に膝をついて臥位になった。なお、後ろ方向に非転倒動作をする際は膝ではなく臀部を床につけた。異常動作時は動画を撮影し、映像から動作開始時刻、異常動作発生時刻、動作終了時刻を記録した。静止動作は転倒した後そのまま静止した状態を想定しているため、転倒動作後、布団の上でそのままの姿勢で静止した。

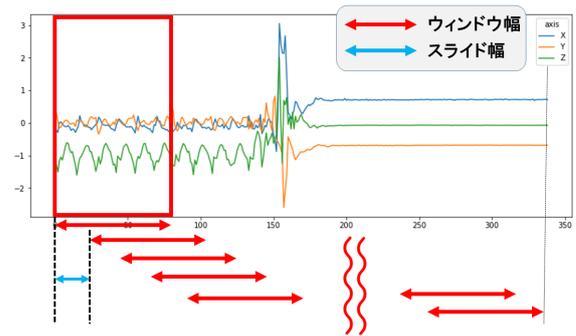


図2: データの分割方法の概要

4.2 学習データの作成

学習させる動作は、歩行動作、転倒動作、非転倒動作、静止状態であり、転倒時の動作方向検出を行うため、歩行動作以外の各動作については、前後左右の4方向について取得した。データは各動作を200回ずつ（歩行動作以外の動作については各方向50回ずつ）取得した。

学習データの長さは、異常動作の動作時間に合わせ、すべての動作について、1データ当たり3秒の長さに統一した。歩行動作については歩行開始時刻から歩行終了時刻の間で3秒間ごとに分割した。異常動作については異常動作発生時刻を基準とし、前1秒間、後2秒間となるよう分割した。静止状態については歩行動作と同様に静止開始時刻から静止終了時刻の間で3秒間ごとに分割した。

4.3 テストデータの作成

テストデータは歩行中に突然転倒動作をし、そのまま静止した際の時系列信号（テストデータA）と、歩行中に突然非転倒動作をし、そのまま静止した際の時系列信号（テストデータB）の2種類とし、それぞれ10回ずつ取得した。データは平坦な道を歩き、途中に設置した布団で異常動作をしてそのまま静止することで取得した。このとき、歩行動作と静止状態については15秒程行った。テストデータは異常動作発生時刻を基準として前5秒間、後8秒間の合計13秒間を使用し、データの分割はスライディングウィンドウ方式で処理を行った（図2参照）。このとき、ウィンドウ幅は学習データと同じ3秒間とし、スライド幅は加速度については5データ分、角速度については2データ分とした。

5 分析

5.1 分類器の作成と評価

学習データにラベル付けを行い、動作の予測を行う分類器を作成したのち、テストデータを用いた動作の識別精度を分析する。分類器の作成には k-Shape 法を用いた。学習データの各動作にそれぞれラベル付けを行った上でクラスタリングを行った。ラベルは歩行、前転倒、後転倒、左転倒、右転倒、前非転倒、後非転倒、左非転倒、右非転倒、前静止、後静止、左静止、右静止の 13 種類とし、このときのクラスタ数は学習データの種類数と同じ 13 に設定した。

学習データを用いてクラスタリングを行った結果を表 1, 2 に示す。数値は各クラスに含まれる動作ごとの学習データ数を示しており、括弧内はクラス内における動作の占有率を示す。占有率が高いほど、当該クラスが示す動作である可能性が高い。

本稿ではこれをもとに、動作、方向ごとの占有率を集計し、最も高いものを当該クラスが示す動作、方向と判別し、占有率を確信度として示した。但し、確信度が 0.5 を超えないものに関しては「判別不可」とした。判別結果を表 3, 4 に示す。

5.2 定常/異常動作識別の手順

テストデータを歩行範囲、異常範囲、静止範囲の 3 つに分類し、精度を検証する (図 3 参照)。範囲の分類方法は以下の計算式で定義した。

$$\text{異常範囲開始} = \frac{\text{異常動作のピーク} - 3 \text{ 秒の長さ}}{\text{スライド幅}} \quad (1)$$

$$\text{静止範囲開始} = \text{異常範囲開始} + \frac{\text{異常動作の範囲}}{\text{スライド幅}} \quad (2)$$

例えばテストデータ A の加速度の場合、異常動作のピークが 155、加速度の 3 秒間のデータ数が 77、スライド幅が 5、異常動作の範囲が 75 であるため、異常範囲と静止範囲の開始ウィンドウ番号は以下ようになる*2。

$$\text{異常範囲開始} : \frac{155 - 77}{5} = 15.6 \quad (3)$$

$$\text{静止範囲開始} : 16 + \frac{75}{5} = 31 \quad (4)$$

したがってそれぞれの範囲におけるウィンドウ番号は、歩行範囲が 0 から 15、異常範囲が 16 から 30、静止範囲が 31 から 53 である。

*2 異常動作のピークおよび範囲は目視で確認した

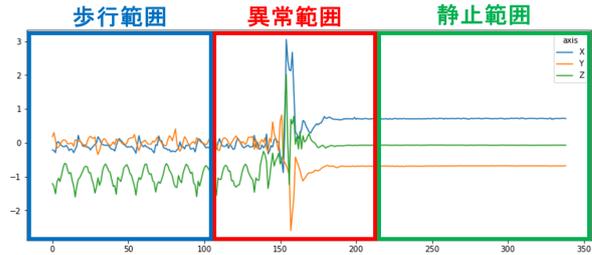


図 3: 範囲の決定のイメージ

5.3 定常/異常動作識別の結果と考察

テストデータを用いた分類器による識別結果 (一部抜粋) を表 5, 6, 7, 8 に示す。左からウィンドウ番号、クラス番号、実際の動作の種類、動作方向、予測による動作の種類、動作方向を示す。クラス番号は分類器によって分類されたクラスを示しており、実際の動作の種類、動作方向は設定した 3 つの範囲をもとに決定した。なお、範囲の境を横線で示す。予測による動作の種類、動作方向は、表 3, 4 をもとにクラス番号が示すものを表す。範囲の設定に対する動作の正解率を表 10 に示す。テストデータ A とテストデータ B それぞれ 10 個のデータを加速度、角速度別で表しており、上の 0 から 9 の数字はテストデータ番号を示している。正解率の求め方は以下のとおりである。

定常/異常動作の正解率

$$= \frac{\text{実際の動作と予測動作が一致しているウィンドウ数}}{\text{すべてのウィンドウ数}} \quad (5)$$

実際の動作と予測動作が一致しているウィンドウ数は表 9 を参照し計算した。その結果、どのデータも正解率平均 0.9 を超える精度だった。テストデータ B の加速度データのみ他に比べて正解率が低いのは、本来「非転倒」と判別しなければならないところを、「転倒」と判断したり、本来静止範囲になければならない「静止」が異常範囲に含まれていることが要因だと考えられる (表 7 参照)。

5.4 転倒/非転倒動作識別の手順

テストデータで設定した異常範囲における、転倒動作と非転倒動作の識別精度を検証する。転倒時の時系列信号では、大きく動いているのが転倒した瞬間だけであることから、転倒/非転倒動作の識別は信号の動きが激しい部分、転倒時の衝撃を対象として行う。本稿では目視にて信号を確認した。転倒時の衝撃が確認できるデー

表 1: 分類器の内訳 (加速度)

クラス	前転倒	後転倒	左転倒	右転倒	前非転倒	後非転倒	左非転倒	右非転倒	歩行	前静止	後静止	左静止	右静止	最大値
0	0(0.000)	13(0.382)	0(0.000)	21(0.618)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.618
1	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	200(0.957)	0(0.000)	0(0.000)	9(0.043)	0(0.000)	0.957
2	1(0.010)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	21(0.219)	17(0.177)	37(0.385)	20(0.208)	0.385
3	0(0.000)	2(0.080)	0(0.000)	0(0.000)	0(0.000)	23(0.920)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.920
4	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	16(0.200)	33(0.413)	1(0.013)	30(0.375)	0.413
5	0(0.000)	13(0.929)	0(0.000)	0(0.000)	0(0.000)	1(0.071)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.929
6	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	46(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000
7	34(0.540)	0(0.000)	0(0.000)	29(0.460)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.540
8	15(0.600)	0(0.000)	0(0.000)	0(0.000)	10(0.400)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.600
9	0(0.000)	15(0.469)	17(0.531)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.531
10	0(0.000)	3(0.077)	33(0.846)	0(0.000)	0(0.000)	0(0.000)	3(0.077)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.846
11	0(0.000)	4(0.148)	0(0.000)	0(0.000)	0(0.000)	23(0.852)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.852
12	0(0.000)	0(0.000)	0(0.000)	0(0.000)	40(0.364)	3(0.027)	1(0.009)	50(0.455)	0(0.000)	13(0.118)	0(0.000)	3(0.027)	0(0.000)	0.455

表 2: 分類器の内訳 (角速度)

クラス	前転倒	後転倒	左転倒	右転倒	前非転倒	後非転倒	左非転倒	右非転倒	歩行	前静止	後静止	左静止	右静止	最大値
0	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	45(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000
1	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	50(0.250)	50(0.250)	50(0.250)	50(0.250)	0.250
2	0(0.000)	20(0.588)	0(0.000)	1(0.029)	0(0.000)	0(0.000)	0(0.000)	13(0.382)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.588
3	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	37(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000
4	14(0.412)	0(0.000)	0(0.000)	0(0.000)	16(0.471)	0(0.000)	2(0.059)	0(0.000)	2(0.059)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.471
5	36(0.900)	0(0.000)	0(0.000)	2(0.050)	2(0.050)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.900
6	0(0.000)	21(0.309)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	47(0.691)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.691
7	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	52(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000
8	0(0.000)	0(0.000)	50(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000
9	0(0.000)	0(0.000)	0(0.000)	0(0.000)	32(0.941)	0(0.000)	1(0.029)	0(0.000)	1(0.029)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.941
10	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	100(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000
11	0(0.000)	9(0.153)	0(0.000)	0(0.000)	0(0.000)	50(0.847)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0.847
12	0(0.000)	0(0.000)	0(0.000)	47(1.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	0(0.000)	1.000

表 3: クラスごとの判別結果 (加速度)

クラス	動作	確信度	方向	確信度
0	転倒	1.000	右	0.618
1	歩行	0.957	判別不可	
2	静止	0.990	判別不可	
3	非転倒	0.920	後	1.000
4	静止	1.000	判別不可	
5	転倒	0.929	後	1.000
6	非転倒	1.000	左	1.000
7	転倒	1.000	前	0.540
8	転倒	0.600	前	1.000
9	転倒	1.000	左	0.531
10	転倒	0.923	左	0.923
11	非転倒	0.852	後	1.000
12	非転倒	0.855	判別不可	

表 4: クラスごとの判別結果 (角速度)

クラス	動作	確信度	方向	確信度
0	歩行	1.000	判別不可	
1	静止	1.000	判別不可	
2	転倒	0.618	後	0.588
3	非転倒	1.000	右	1.000
4	非転倒	0.529	前	0.882
5	転倒	0.950	前	0.950
6	非転倒	0.691	左	0.691
7	歩行	1.000	判別不可	
8	転倒	1.000	左	1.000
9	非転倒	0.971	前	0.941
10	歩行	1.000	判別不可	
11	非転倒	0.847	後	1.000
12	転倒	1.000	右	1.000

タの範囲は加速度が約 20 データ、角速度が約 8 データだったため、スライド幅で除算するとどちらも 4 となる。したがって異常範囲における最初の 4 ウィンドウ分を識別の対象とした。識別対象のウィンドウにおける動作の正解率を以下のように算出した。

転倒/非転倒動作の正解率

$$= \frac{\text{実際の動作と予測動作が一致しているウィンドウ数}}{\text{識別対象とするウィンドウ数}} \quad (6)$$

実際の動作と予測動作が一致しているウィンドウ数について、テストデータ A の場合は転倒動作と判別されているもの、テストデータ B の場合は非転倒動作と判別されているものを対象とした。

5.5 転倒/非転倒動作識別の結果と考察

算出した正解率を表 11 に示す。両テストデータとも、加速度においては正解率が高かったがテストデータ A の角速度においては、正解がない場合もあり、平均 0.375

表 5: 分類器による識別結果 (テストデータ A 加速度)

ウィンドウ	クラス	実際		予測	
		動作	方向	動作	方向
0	1	歩行	なし	歩行	判別不可
1	1	歩行	なし	歩行	判別不可
⋮	⋮	⋮	⋮	⋮	⋮
15	1	歩行	なし	歩行	判別不可
16	7	転倒	前	転倒	前
17	7	転倒	前	転倒	前
18	7	転倒	前	転倒	前
19	7	転倒	前	転倒	前
20	7	転倒	前	転倒	前
21	7	転倒	前	転倒	前
⋮	⋮	⋮	⋮	⋮	⋮
26	7	転倒	前	転倒	前
27	7	転倒	前	転倒	前
28	7	転倒	前	転倒	前
29	7	転倒	前	転倒	前
30	7	転倒	前	転倒	前
31	2	静止	前	静止	判別不可
32	2	静止	前	静止	判別不可
⋮	⋮	⋮	⋮	⋮	⋮
53	2	静止	前	静止	判別不可

表 7: 分類器による識別結果 (テストデータ B 加速度)

ウィンドウ	クラス	実際		予測	
		動作	方向	動作	方向
0	1	歩行	なし	歩行	判別不可
1	1	歩行	なし	歩行	判別不可
⋮	⋮	⋮	⋮	⋮	⋮
15	1	歩行	なし	歩行	判別不可
16	12	非転倒	前	非転倒	判別不可
17	12	非転倒	前	非転倒	判別不可
18	12	非転倒	前	非転倒	判別不可
19	12	非転倒	前	非転倒	判別不可
20	12	非転倒	前	非転倒	判別不可
21	8	非転倒	前	転倒	前
⋮	⋮	⋮	⋮	⋮	⋮
26	8	非転倒	前	転倒	前
27	2	非転倒	前	静止	判別不可
28	2	非転倒	前	静止	判別不可
29	2	非転倒	前	静止	判別不可
30	2	非転倒	前	静止	判別不可
31	2	静止	前	静止	判別不可
32	2	静止	前	静止	判別不可
⋮	⋮	⋮	⋮	⋮	⋮
53	2	静止	前	静止	判別不可

表 6: 分類器による識別結果 (テストデータ A 角速度)

ウィンドウ	クラス	実際		予測	
		動作	方向	動作	方向
0	10	歩行	なし	歩行	判別不可
1	10	歩行	なし	歩行	判別不可
2	10	歩行	なし	歩行	判別不可
3	0	歩行	なし	歩行	判別不可
4	7	歩行	なし	歩行	判別不可
5	7	歩行	なし	歩行	判別不可
6	10	歩行	なし	歩行	判別不可
7	10	歩行	なし	歩行	判別不可
8	10	歩行	なし	歩行	判別不可
9	0	歩行	なし	歩行	判別不可
10	0	歩行	なし	歩行	判別不可
11	7	歩行	なし	歩行	判別不可
12	7	歩行	なし	歩行	判別不可
13	10	歩行	なし	歩行	判別不可
14	10	歩行	なし	歩行	判別不可
15	10	歩行	なし	歩行	判別不可
16	0	歩行	なし	歩行	判別不可
17	0	歩行	なし	歩行	判別不可
18	7	歩行	なし	歩行	判別不可
19	9	転倒	前	非転倒	前
20	9	転倒	前	非転倒	前
21	5	転倒	前	転倒	前
22	5	転倒	前	転倒	前
23	5	転倒	前	転倒	前
⋮	⋮	⋮	⋮	⋮	⋮
37	5	転倒	前	転倒	前
38	1	静止	前	静止	判別不可
39	10	静止	前	歩行	判別不可
40	1	静止	前	静止	判別不可
⋮	⋮	⋮	⋮	⋮	⋮
66	1	静止	前	静止	判別不可

表 8: 分類器による識別結果 (テストデータ B 角速度)

ウィンドウ	クラス	実際		予測	
		動作	方向	動作	方向
0	7	歩行	なし	歩行	判別不可
1	7	歩行	なし	歩行	判別不可
2	10	歩行	なし	歩行	判別不可
3	10	歩行	なし	歩行	判別不可
4	10	歩行	なし	歩行	判別不可
5	0	歩行	なし	歩行	判別不可
6	7	歩行	なし	歩行	判別不可
7	7	歩行	なし	歩行	判別不可
8	7	歩行	なし	歩行	判別不可
9	10	歩行	なし	歩行	判別不可
10	10	歩行	なし	歩行	判別不可
11	0	歩行	なし	歩行	判別不可
12	0	歩行	なし	歩行	判別不可
13	7	歩行	なし	歩行	判別不可
14	7	歩行	なし	歩行	判別不可
15	7	歩行	なし	歩行	判別不可
16	10	歩行	なし	歩行	判別不可
17	10	歩行	なし	歩行	判別不可
18	9	非転倒	前	非転倒	前
19	9	非転倒	前	非転倒	前
20	9	非転倒	前	非転倒	前
21	4	非転倒	前	非転倒	前
22	4	非転倒	前	非転倒	前
⋮	⋮	⋮	⋮	⋮	⋮
36	4	非転倒	前	非転倒	前
37	1	静止	前	静止	判別不可
38	1	静止	前	静止	判別不可
39	1	静止	前	静止	判別不可
⋮	⋮	⋮	⋮	⋮	⋮
65	1	静止	前	静止	判別不可
66	1	静止	前	静止	判別不可

と識別精度が低かった。テストデータ A の角速度における識別対象範囲の 4 ウィンドウ分で判別された動作の内訳を表 12 に示す。転倒と判別されれば正しく識別できていることになるが、転倒以外の動作で誤判別されている動作の割合は「非転倒動作」が 0.575 を占めている。本来歩行範囲に含まれるはずの「歩行動作」の誤判別が 0.050 と低い割合であることから、テストデータ A の角速度において識別精度が低かったのは「転倒動作」と判別すべきところを「非転倒動作」と誤って判別している

ことが原因だといえる。

以上の結果から、転倒/非転倒動作の判別には加速度の時系列信号を用いたほうが識別精度が高いことが確認された。

5.6 動作方向識別の手順

異常動作発生時の動作方向の識別精度を検証する。対象とするデータは、加速度、角速度のテストデータ A、

表 9: 定常/異常動作正解率算出法

		実際の動作		
		歩行範囲	異常範囲	静止範囲
予測動作	歩行動作	一致	不一致	不一致
	異常動作 ¹	不一致	一致	不一致
	静止状態	不一致	不一致	一致

¹ テストデータ A は転倒動作,
 テストデータ B は非転倒動作の齟齬を確認。

Bにおける異常範囲、静止範囲とする。今回用いたテストデータはすべて前方向への動作をしているため、前方向と判別された結果を正解とし、判別結果から動作方向の正解率を算出する。正解率の算出方法は以下のとおりである。

動作方向の正解率

$$= \frac{\text{実際と予測の動作方向が一致しているウィンドウ数}}{\text{対象範囲のウィンドウ数}} \quad (7)$$

5.7 動作方向識別の結果と考察

分類器によるクラスタリング結果（表 1, 2 参照）で、静止動作の方向判別はできなかった。静止状態の時系列信号のみで身体の状態を推定することは困難だった。異常範囲において算出した正解率を表 13 に示す。両テストデータとも、角速度においては正解率が高かったがテストデータ B の加速度においては、平均 0.407 と識別精度が低かった。テストデータ B の加速度の異常範囲内において判別された動作方向の内訳を表 14 に示す。前以外の方向は判別されておらず、「判別不可」となったものが平均 0.593 を占めている。ここで、「判別不可」となった動作の内訳を表 15 に示す。「非転倒動作」と判別できているものの、動作方向の識別に対する確信度が低く「判別不可」となってしまったものが 0.472 を占め、本来静止範囲に含まれるはずの「静止状態」が 0.517 を占めていた。したがってテストデータ B の加速度において動作方向の正解率が低かった要因は各範囲の設定が正しくできていなかったことがあげられるが、動作の判別は正しくできたのに対し動作方向の判別ができなかったことを考慮すると、加速度を用いて動作方向の識別をすると識別精度が低くなると考えられる。したがって動作方向の識別を行う際には角速度の時系列信号を用いたほうが高い精度で識別できることを確認した。

6 おわりに

本稿では機械学習の k-Shape 法を用いて動作の予測を行い、その識別精度を検証した。独居高齢者は一人で

暮らしているため、転倒時に助けを必要とする場合でも周囲の人から発見されるまで時間がかかる。本研究は独居高齢者の転倒を検出し、他者への救助要請を行うことを目的とし、ウェアラブルセンサを用いた転倒動作検出を試みた。提案手法では学習データを用いて分類器を作成した後、テストデータを用いて分類器による動作の識別精度を検証した。その結果、転倒/非転倒動作の識別には加速度の時系列信号を、動作方向の識別には角速度の時系列信号を用いることで、精度高く動作の識別ができることが確認できた。今後は、表 3 のラベル 12 のような、動作の識別はできたが、動作方向の識別ができなかった判別結果をなくすため、学習データに用いるデータ数を増やして分類器の精度を上げる必要がある。また表 7 のウィンドウ番号 27 から 30 のように、本来当てる範囲と異なる範囲に誤設定してしまっていることから、範囲の設定方法についても改めて考慮する必要がある。本稿で扱った動作は日常生活の中で起こりうる動作のごく一部であるため、扱う動作の種類を増やし、心電、心拍数や、深部温度データを用いたセンサ装着者の身体状態のモニタリングも同時に行えるシステムの実現を目指す。

参考文献

- [1] De, D., Bharti, P., Das, S. K. and Chellappan, S.: Multimodal Wearable Sensing for Fine-Grained Activity Recognition in Healthcare, *IEEE*, Vol. 19, No. 5, pp. 26–35 (2015).
- [2] Farrahi, V., Niemelä, M., Kangas, M., Korpelainen, R. and Jämsä, T.: Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches, *Gait & Posture*, Vol. 68, pp. 285–299 (2019).
- [3] Kwapisz, J. R., Weiss, G. M. and Moore, S. A.: Activity recognition using cell phone accelerometers, *ACM SIGKDD Explorations Newsletter*, Vol. 12, No. 2, pp. 74–82 (2010).
- [4] 磯和之, 白井蘭隆也: テレワーク環境におけるアンビエントセンサを用いた作業者の状態推定に関する実験, 情報処理学会研究報告, Vol. 2018-GN-104, No. 19, pp. 1–7 (2018).
- [5] 小川裕也, 内藤克浩: 放射温度センサを用いた温度分布による転倒検出, 情報処理学会研究報告, Vol. 2019-CDS-26, No. 4, pp. 1–6 (2019).
- [6] 高橋雄太, 音田恭宏, 藤本まなと, 荒川豊: センサ装

表 10: 定常/異常動作識別正解率

	0	1	2	3	4	5	6	7	8	9	平均	最大値	最小値
加速度 (A 転倒)	0.926	0.944	1.000	1.000	1.000	0.963	0.963	1.000	0.944	1.000	0.974	1.000	0.926
加速度 (B 非転倒)	0.907	0.926	0.926	0.926	0.907	0.907	0.870	0.926	0.889	0.852	0.904	0.926	0.852
角速度 (A 転倒)	0.985	0.985	1.000	0.970	0.970	0.985	0.940	0.985	0.955	0.985	0.976	1.000	0.940
角速度 (B 非転倒)	1.000	1.000	1.000	1.000	1.000	0.970	0.985	1.000	0.970	1.000	0.993	1.000	0.970

表 11: 転倒/非転倒動作識別正解率

	0	1	2	3	4	5	6	7	8	9	平均	最大値	最小値
加速度 (A 転倒)	1.000	1.000	1.000	1.000	1.000	0.750	1.000	1.000	1.000	1.000	0.975	1.000	0.750
加速度 (B 非転倒)	1.000	1.000	1.000	1.000	1.000	0.750	0.750	1.000	1.000	0.750	0.925	1.000	0.750
角速度 (A 転倒)	0.750	0.500	0.000	0.750	0.500	0.000	0.500	0.000	0.750	0.000	0.375	0.750	0.000
角速度 (B 非転倒)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.750	1.000	0.975	1.000	0.750

表 12: テストデータ A (角速度) における転倒/非転倒動作識別対象範囲の動作内訳

	0	1	2	3	4	5	6	7	8	9	平均
転倒	0.750	0.500	0.000	0.750	0.500	0.000	0.500	0.000	0.750	0.000	0.375
非転倒	0.000	0.500	1.000	0.250	0.500	1.000	0.250	1.000	0.250	1.000	0.575
歩行	0.250	0.000	0.000	0.000	0.000	0.000	0.250	0.000	0.000	0.000	0.050
静止	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

表 13: 動作方向正解率

	0	1	2	3	4	5	6	7	8	9	平均	最大値	最小値
加速度 (A 転倒)	0.733	0.800	1.000	1.000	1.000	0.800	0.933	1.000	0.800	1.000	0.907	1.000	0.733
加速度 (B 非転倒)	0.400	0.400	0.400	0.467	0.467	0.400	0.400	0.400	0.333	0.400	0.407	0.467	0.333
角速度 (A 転倒)	0.800	1.000	1.000	1.000	1.000	0.933	0.933	1.000	1.000	0.933	0.960	1.000	0.800
角速度 (B 非転倒)	1.000	1.000	1.000	1.000	1.000	0.933	1.000	1.000	0.933	1.000	0.987	1.000	0.933

表 14: テストデータ B (加速度) における異常範囲内の動作方向内訳

	0	1	2	3	4	5	6	7	8	9	平均
前	0.400	0.400	0.400	0.467	0.467	0.400	0.400	0.400	0.333	0.400	0.407
後	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
左	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
右	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
判別不可	0.600	0.600	0.600	0.533	0.533	0.600	0.600	0.600	0.667	0.600	0.593

表 15: テストデータ B (加速度) の異常範囲内において動作方向が判別不可だった動作内訳

	0	1	2	3	4	5	6	7	8	9	平均
転倒	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
非転倒	0.444	0.556	0.556	0.500	0.500	0.444	0.333	0.556	0.500	0.333	0.472
歩行	0.000	0.000	0.000	0.000	0.000	0.111	0.000	0.000	0.000	0.000	0.011
静止	0.556	0.444	0.444	0.500	0.500	0.444	0.667	0.444	0.500	0.667	0.517

着杖を介した歩行動作検出手法の提案, 情報処理学会論文誌コンシューマ・デバイス&システム (CDS), Vol. 8, No. 2, pp. 43-55 (2018).

- [7] 橋本浩二, 吉野春香, モシニヤガワシリー: ドップラーセンサと機械学習による人物転倒検出システムの開発, 組込みシステムシンポジウム 2019 論文集, pp. 13-16 (2019).

帰納論理プログラミングを用いた不動産間取り図からの選好抽出

Preference Learning from Real Estate Floor Plans using Inductive Learning of Answer Set Programs

橋本 琉¹ 尾崎 知伸^{1*}
Ryu Hashimoto¹ Tomonobu Ozaki¹

¹ 日本大学 文理学部

¹ College of Humanities and Sciences, Nihon University

Abstract: In this paper, we discuss preference learning for real estate floor plans. Specifically, we use Answer Set Programs to represent the connections among rooms as well as some properties of each room such as size and shape. Then, weak constraints representing the preferences on some purposes will be extracted using inductive preference learning of answer set programs. The usefulness of our proposal are assessed using different purposes among nine rental properties.

1 はじめに

不動産物件の選好基準には、各部屋の大きさや形、窓や収納の数など部屋に関する属性に加え、部屋同士の繋がりや配置などの多様な要因が考えられ、またそれらは物件の使用用途に大きく依存する。使用用途に応じた具体的な選好基準を優先順位と共に明示化することは、借り手に対する物件選びの支援に加え、貸し手に対する賃料設定・物件アピールの根拠の提供や、物件設計者に対する新たな検討要素の提示に繋がり、その影響は大きいと考えられる。

これらのことを背景に、本研究では、部屋属性と部屋配置に着目し、使用用途に関する物件ランキングからその明示的な選好基準を獲得することを考える。具体的には、論理プログラム的一种である解集合プログラミング (Answer Set Programming, ASP) [1, 2] を用いて部屋属性と部屋配置を表現するとともに、ASP 上の帰納推論・選好学習の枠組みである ILASP (Inductive Learning of ASP) [3, 4] を利用し、選好基準を表す論理的規則 (弱制約) 群の抽出を試みる。

本論文の構成は以下の通りである。2 章で関連研究について言及する。また 3 章で ILASP による選好学習について説明する。これらを基に、4 章で不動産物件に関する選好学習の評価実験を行い、最後に 5 章でまとめと今後の課題を述べる。

2 関連研究

これまでに種々の要因に着目した不動産賃貸物件の分析が行われている。特に近年では、部屋配置 (間取り) に着目した研究も多く、不動産間取り図から得られる情報を基に各物件を部屋を頂点、その繋がりを辺とするグラフに変換し、グラフ理論やグラフマイニング技術を用いた分析が行われている。例えば文献 [5] では平面構成の分析が、文献 [6] や [7] では賃料分析がそれぞれ行われている。また、賃料への影響が大きい間取りの積極的な抽出 [8, 9] や説明可能 AI 技術を用いた賃料推定根拠の提示 [10] に加え、文献 [11] では、グラフ構造と住み心地の関連についての分析が展開されている。これらの関連研究に対し、本研究の中心課題は論理表現された間取り図を対象とした選好学習であり、表現・分析手法の面で大きく異なっている。

これまでに、論理プログラムを用いた選好学習に関し、いくつかの応用が報告されている。文献 [12] では、SVM や決定木に加え、帰納論理プログラミングを用い、分類問題の枠組みを用いた選好学習について議論している。また文献 [13] では、学習済み SVM モデルをオラクルとして利用することで学習データを構築するとともに、ILASP による選好学習の枠組みを利用して、局所的な選好に関する説明及びモデルの大域的な説明を生成している。いずれの研究も、分析対象自身は命題論理で表現されるが、本研究では、述語論理を用いて表現されるより複雑な問題を対象としており、その点で大きく異なっている。

*連絡先：日本大学文理学部情報科学科
〒156-8550 東京都世田谷区桜上水 3-25-40
E-mail: ozaki.tomonobu@nihon-u.ac.jp

3 ILASP による選好学習

選好の抽出に際し、部屋の繋がりだけでなく、どのような部屋属性をどれ程度含んでいるかも考慮することを考えた場合、多くの既存研究で採用されているラベル付きグラフによる表現は必ずしも十分ではない。そこで本研究では、述語論理に基づき、また量の概念も扱うことのできる解集合プログラミング [1, 2] を表現言語として採用する。また分析手法として、帰納論理プログラミングの近年の発展 [14] の一つである、解集合プログラムを対象とした帰納推論・選好学習の枠組みである ILASP [3, 4] を採用する。以下では、本研究に直接関係のある部分を中心に、ILASP による選好学習の枠組みを説明する。

ILASP への入力は大きく、事例集合 $E = E^+ \cup E^-$ 、選好対 (順序事例) 集合 O 、背景知識 B 、仮説空間 S_M に分けられる。各事例 $e = (e^{inc}, e^{exc}, e^{ctx}) \in E$ は、部分解釈に相当する 2 つの基礎アトムの集合 e^{inc}, e^{exc} と、 e に固有の背景知識としての解集合プログラム e^{ctx} から構成される。また、選好対 (順序事例) $o = (e_1, e_2) \in O$ は 2 つの事例 e_1, e_2 から構成され、 e_1 より e_2 を好む ($e_1 \succ e_2$) という選好を表現する。解集合プログラム B は事例に共通の背景知識であり、また仮説空間 S_M は、モード宣言 M 等を用いて構築される、弱制約を含む解集合プログラムの集合である。ここで弱制約とは、 $b_1, \dots, b_n, c_1, \dots, c_m$ を原子式、 t_1, \dots, t_o を項、 w と l をそれぞれ重みとレベルを表す項としたとき

$$:\sim b_1, \dots, b_n, \text{not } c_1, \dots, \text{not } c_m. \\ [w@l, t_1, \dots, t_o]$$

の形で表現されるルールであり、モデル間の選好基準を表す。より形式的には、プログラム P の安定モデルの集合を $AS(P)$ としたとき、 P と $A \in AP(P)$ に対し、レベル l 毎に

$$P_A^l = \sum_{(w,l,t_1,\dots,t_o) \in \text{weak}(P,A)} w \quad \text{where} \\ \text{weak}(P,A) = \{(w,l,t_1,\dots,t_o) \mid \\ A \text{ は } b_1, \dots, b_n, \text{not } c_1, \dots, \text{not } c_m \text{ を満たす}\}$$

なる値 P_A^l を計算し、 l を大きさを第一基準、 P_A^l の小ささを第二基準として、モデル間の優先順位を決定する。また $P_{A_1}^{l_1}$ より $P_{A_2}^{l_2}$ の優先順位が高いとき、それを $P_{A_1}^{l_1} \succ P_{A_2}^{l_2}$ と表現する。

以上の準備のもと、ILASP では $E = E^+ \cup E^-$ 、 O 、 B 、 S_M を入力とし、以下の 3 条件を満たす (リテラル総数に関して) 最小の仮説 $H \subseteq S_M$ を導出する。詳細は、原著論文 [3, 4] を参照されたい。

1. $\forall e \in E^+, \exists A \in AS(B \cup e^{ctx} \cup H)$
 $[e^{inc} \subseteq A, e^{exc} \cap A = \emptyset]$
2. $\forall e \in E^-, \nexists A \in AS(B \cup e^{ctx} \cup H)$
 $[e^{inc} \subseteq A, e^{exc} \cap A = \emptyset]$

3. $\forall (e_1, e_2) \in O,$
 $\exists A_1 \in AS(B \cup e_1^{ctx} \cup H), A_2 \in AS(B \cup e_2^{ctx} \cup H)$
 $[P_{A_1}^{l_1} \succ P_{A_2}^{l_2}]$

なお本研究では選好学習にのみ焦点を当てており、 E^- 及び $e \in E^+$ における e^{inc}, e^{exc} は空集合であり、また S_M は弱制約のみを含む。従って実際には、上記の第 3 条件のみを考慮した弱制約の抽出 (選択) が行われることになる。

4 間取り図の論理表現と選好の抽出

4.1 対象データと述語表現

本研究では、国立情報学研究所の IDR データセット提供サービスにより株式会社 LIFULL から提供を受けた「LIFULL HOME'S データセット」を利用した。データセットから、東急田園都市線の溝の口駅を最寄り駅とする、賃料や部屋面積に大きな差のない 2LDK, 3LDK の物件 9 件を選定し、実験に使用した。

本研究で採用した論理表現 (述語) を図 1 に示す。なお、物件情報の論理表現への変換は、著者らが手作業で行った。

まず、部屋属性に関する述語について説明する。述語 `room_type` は、各部屋の種別を表すものである。本研究では 9 つの部屋種別 (玄関、廊下、風呂、洗面所、トイレ、洋室、和室、LDK、ベランダ) を準備した。また `room_shape` は部屋の概形を表すもので、各部屋を (ほぼ) 正方形、長方形、それ以外の 3 種に分類する。`has_window`, `has_closet`, `has_door` の 3 述語は、各部屋に付随する設備数を表すものであり、該当する設備を 1 つ以上含む部屋に対して準備する。また `has_wm_space`, `has_cooking_space` も同様、該当する部屋に対してのみ準備している。

次に、部屋配置に関する述語について説明する。`link` は部屋の繋がりを表す最も基本的な述語であり、部屋種別を表す `room_type` 述語と合わせることで、従来のラベル付きグラフによる表現を再現している。一方、`hop_to_entrance` は玄関から各部屋までの移動数に相当し、玄関と部屋間の最短経路長に相当する。また `corridor_to_rooms` は、廊下に対して定義される連結された部屋数を表す述語である。`independent_room` は、個室利用可能な部屋、すなわち次数 1 の洋室・和室を意味する。`hop_to_entrance` 以降の 3 述語は、例えば

```
corridor_to_rooms(R, NR) :-
    room_type(R, corridor), NR = { link(R, _) }.
independent_room(R) :-
    room_type(R, jbed), 1 = { link(R, _) }.
```

部屋属性に関する述語	
●	room_type(R, T): 部屋 R のタイプは T である
●	room_size(R, A): 部屋 R の面積は A である
●	room_shap(R, S): 部屋 R の形は S である
●	has_window(R, NW): 部屋 R は NW 個の窓を持つ
●	has_closet(R, NC): 部屋 R は NC 個の収納を持つ
●	has_door(R, ND): 部屋 R は ND 個のドアを持つ
●	has_wm_space(R): 部屋 R には洗濯機置き場がある
●	has_cooking_space(R): 部屋 R には調理空間がある
部屋配置に関する述語	
●	link(R1, R2): 部屋 R1 と部屋 R2 は繋がっている
●	hop_to_entrance(R, NH): 玄関から部屋 R までは NH ホップである
●	corridor_to_rooms(R, NR): 廊下 R は NR の部屋とつながっている
●	independent_room(R): 部屋 R は個室としての利用可能である

図 1: 間取り図に対する述語表現

などのルールを用いて表現することが可能である。しかし今回は、計算時間を考慮し、事前に計算した上でファクトとしてシステムへと与えている。

4.2 選好対の設定

本研究では、対象 9 物件に対する下記のランキングから選好対を設定した。

1. 賃料に基づくランキング (賃料の降順)
2. 単独での利用 (一人暮らし) を想定した場合のランキング
3. 友人 2 名での利用 (シェアハウス) を想定した場合のランキング

賃料以外のランキングは、大学生 6 名によるアンケート調査より決定している。表 1 に 9 物件のランキング

表 1: 物件のランキング

物件	賃料	一人暮らし	シェアハウス
物件 1	1	5	5
物件 2	2	4	3
物件 3	3	6	7
物件 4	4	7	6
物件 5	5	3	4
物件 6	6	8	9
物件 7	7	2	1
物件 8	7	1	2
物件 9	9	9	8

を示す。表から、使用目的によってランキングが大きく変わっていることが分かる。

各ランキングからの選好対の生成は、第 i 位と $i+1$ 位の間に顕著な差が認められる場合に選好対

ランク i 位の物件 \succ ランク $i+1$ 位の物件

を生成することを基本とし、いくつかの例外を追加・削除する形で行った。結果として、賃料ランキング、一人暮らしランキング、シェアハウスランキングのそれぞれで、9 件、7 件、8 件の選好対が得られ、それらを教師情報として利用した。

4.3 結果と考察

各ランキングに対して得られた選好基準 (弱制約) を図 2 に示す。ここで弱制約は、 $[w@l, t_1, \dots, t_o]$ の l の大きさを第一の基準に、項 t_1, \dots, t_o の具体化に対して与えられる w の総和の小ささを第二の基準にすることに注意されたい。なお今回の実験では、現実的な時間で解を獲得するために、各弱制約のリテラル数および変数の数の上限を共に 3 に設定している。加えて、存在しない部屋配置 (「バルコニーとトイレが繋がっている」など) や部屋属性 (「バルコニーに窓がある」など) を排除するとともに、より意味に立ち入った制限 (「部屋属性を伴わないリンク情報は認めない」など) を導入することで、適切な範囲に仮説空間を限定し、計算爆発を回避している。

図 2 に示す通り、賃料に関するランキングからは「窓がある部屋の合計面積が大きいほど家賃が安い」という選好基準が獲得された。この選好基準は、“窓” や “大きさ” という部屋属性を中心に構成されており、単純なラベル付きグラフでは表現することのできない基準となっている。

また、一人暮らしに関するランキングからは、第一の基準として「ドア総数が多い方が好ましい」、第二の基準として「個室利用できる部屋の合計面積が大きい

- 賃料ランキング
:~ has_window(R,W), size_of_room(R,Size).
[Size@1, R, W, Size]
- 一人暮らしランキング
:~ hops_to_entrance(R,E), has_door(R,Door).
[-Door@2, R, Door, E]
:~ independent_room(R), size_of_room(R,Size).
[-Size@1, R, Size]
- シェアハウスランキング
:~ has_window(R,W), has_door(R,Door).
[-Door@1, R, Door, W]

図 2: 得られた弱制約

いほど好ましい」という2つの選好基準が獲得された。優先順位を伴う複数の選好基準の抽出に加え、第二の選好基準にみられる部屋配置（個室利用）と部屋属性（面積）を併用するルールは、他の関連研究では扱うことが困難なものであり、本アプローチの有用性を示すものであると考えている。

一方、シェアハウスに関するランキングからは「ドアと窓がある部屋の総ドア数が多いほど好ましい」という選好基準が獲得された。賃料ランキングの場合と同様、この選好基準は複数の部屋属性を伴うものであり、本研究の狙いの一つである、部屋属性への着目が適切に反映されたものであると考えている。

以上示した通り、今回の実験では、3種のランキングそれぞれにおいて異なる選好基準が抽出された。これらの結果は初期的なものではあるが、使用用途に応じた多様な選好基準の表現・獲得に関する本アプローチの有効性を示すものであると考えている。

5 まとめ

本研究では、解集合プログラミング上での選好学習の枠組みを利用し、不動産賃貸物件に関する種々の選好基準を明示的な形で抽出することを試みた。その結果、各ランキングに対してそれぞれ異なる規則群の抽出に成功した。

今後の課題として、大規模データと多様なランキングを用いた追加実験が挙げられる。また、ランキングの上位・下位を識別する分類問題の枠組みを通じて得られる選好との比較も検討している。加えて、利用述語を再検討することで、説得性・納得性の高い選好基

準を抽出することも重要な課題の一つである。

謝辞: 本研究では、国立情報学研究所の IDR データセット提供サービスにより株式会社 LIFULL から提供を受けた「LIFULL HOME'S データセット」を利用した。

参考文献

- [1] 坂間 千秋, 井上 克巳: 解集合プログラミング, 人工知能学会誌, 25(3):368-378, 2010.
- [2] 井上 克巳, 坂間 千秋: 論理プログラミングから解集合プログラミングへ, コンピュータソフトウェア, 25(3):20-32, 2008.
- [3] M. Law, A. Russo, and K. Broda: The ILASP system for Inductive Learning of Answer Set Programs, arXiv:2005.00904 [cs.AI], 2020.
- [4] M. Law, A. Russo, and K. Broda: Inductive learning of answer set programs, *Proc. of the 14th European Workshop on Logics in Artificial Intelligence*, pp.311-325, 2014.
- [5] 花里 俊廣, 平野 雄介, 佐々木 誠: 首都圏で供給される民間分譲マンション 100m² 超住戸の隣接グラフによる分析, 日本建築学会計画系論文集 (591), pp.9-16, 2005.
- [6] 瀧澤 重志, 吉田 一馬, 加藤 直樹: グラフマイニングを用いた室配置を考慮した賃料分析: 京都市郊外の 3LDK を中心とした賃貸マンションを対象として, 日本建築学会環境系論文集 (623):139-146, 2008.
- [7] 尾崎 知伸, 小黒 淳斗: 頻出部分グラフを用いた賃料分析, 人工知能学会第 111 回知識ベースシステム研究会, pp.13-16, 2017.
- [8] 長谷川 優也, 尾崎 知伸: 部屋配置とその出現数に着目した二段階賃料推定, 人工知能学会 第 21 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.80-87, 2019.
- [9] Tomonobu Ozaki: Extraction of Characteristic Subgraph Patterns with Support Threshold from Databases of Floor Plans, *Proc of the 2019 Seventh International Symposium on Computing and Networking*, pp.197-203, 2019.
- [10] 菱沼 大輝, 尾崎 知伸: 傾向スコア分析と説明可能 AI 技術を用いた賃料決定に関する重要部屋配置の特定, 第 206 回情報処理学会 知能システム研究会, 2021.
- [11] 橋原 太郎, 汪 雪, 山崎 俊彦: 不動産間取図の魅力度データを使った「住み心地」とグラフ構造の分析, 日本人工知能学会第 35 回大会, 4F3-GS-10n-04, 2021.
- [12] N. N. Qomariyah and D. Kazakov: Learning Binary preference Relations, *Proc. of the 4th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems*, pp.30-34, 2017.

- [13] F. A. D' Asaro, M. Spezialetti, L. Raggioli, S. Rossi :
Towards an Inductive Logic Programming Approach
for Explaining Black-Box Preference Learning Sys-
tems, *Proc. of the 17th International Conference on
Principles of Knowledge Representation and Reason-
ing — Special Session on KR and Machine Learning*,
pp.855–859, 2020.
- [14] A. Cropper, S. Dumančić, S. H. Muggleton : Turn-
ing 30: New Ideas in Inductive Logic Programming,
*Proc. of the 29th International Joint Conference on
Artificial Intelligence*, Survey track, pp.4833-4839,
2020.
- [15] 株式会社 LIFULL : LIFULL HOME'S データセット,
国立情報学研究所情報学研究データリポジトリ (データ
セット) <https://doi.org/10.32130/idr.6.0>, 2015.

スムーズなテキスト分析を支援する音声ナビゲーション

Voice Navigation to Support Smooth Text Mining

佐藤 允哉^{1*} 砂山 渡²
Masaya Sato¹ Wataru Sunayama²

¹ 滋賀県立大学大学院 工学研究科

¹ Graduate School of Engineering, The University of Shiga Prefecture

² 滋賀県立大学 工学部

² School of Engineering, The University of Shiga Prefecture

Abstract: In recent years, the sophistication of information and communication technologies has led to a continuous accumulation of data, but the training of data scientists to perform data mining has not kept pace. In addition, to learn the flow of data mining, it is necessary to understand the mining procedure, but it is currently difficult to learn this by oneself. In this study, we propose a voice navigation system using TETDM, a text mining tool, to assist those who are not familiar with data mining in smoothly executing the mining procedure, from determining the purpose of the mining to obtaining knowledge. The experimental results show that the proposed voice navigation system is effective in supporting the smooth execution of the text mining procedure.

1 はじめに

近年、IT技術の発展によってデータのやり取りは紙媒体ではなくデジタルデータとして取り扱われることが多くなっている。その影響でビッグデータと呼ばれる膨大な量のデータが蓄積され続けている。それらのデータの中には、顧客アンケートの結果や、SNSに書き込まれた製品の評価といった企業戦略や商品開発に使えるような重要なデータも含まれている。

また、これらのデータには数値データやテキストデータと様々な種類のデータが含まれており、分析を行う際にはそれらのデータの種類に応じた分析手法をとる必要がある。中でも、テキストデータの分析は、膨大な量のテキストデータを読まなくてはならず、尚且つテキストデータの種類によっても分析の手法が変わってくるため解析の目的にあったデータ取得方法を考える必要がある [1]。

これまでも、テキスト分析のツールの使い方や分析の手順のチュートリアルによって、テキスト分析の基本的な知識を身につけさせるシステムが提案されてきた。しかしながら、チュートリアルだけでは、テキスト分析の手順やツールの操作方法について理解しにくいことが存在し、それらのつまづきによってテキスト分析がスムーズに行われない可能性がある。

そこで本研究では、テキスト分析を行うための統合環境である TETDM に新たな機能として、音声ナビゲーションによるテキスト分析支援を行う機能を実装する。この音声ナビゲーション機能に従いながらテキスト分析を進めることで、ツールの使い方や次に行う動作が理解しやすくなり、これらによってテキストデータから有益な情報をスムーズに取り出すことを目的としている。

2 関連研究

2.1 データサイエンスのスキル獲得支援の研究

データサイエンスのスキル獲得を支援する研究 [2] がある。この研究では、初心者ユーザを対象に、テキスト分析スキルの獲得を目的としたチュートリアルシステムを開発し、それらによって基礎的なスキルから応用的なスキルまで順にテキスト分析課題を解かせることで、ユーザの一定のテキスト分析のスキル向上により支援を行った。しかしながら、これらの研究ではチュートリアル用としてあらかじめ用意されたテキストで行うので、ユーザが分析したいテキストとは同じように進められない可能性がある。そこで本研究では、実際に分析を行いながら音声ナビゲーションによるテキスト分析支援を行うことで、手順やツールの操作方法を

*連絡先：滋賀県立大学大学院工学研究科 電子システム工学専攻
佐藤 允哉
〒 522-8533 滋賀県彦根市八坂町 2500
E-mail: os23msato@ec.usp.ac.jp

習得してもらうことを目的とした。

2.2 音声ナビゲーションによる操作支援

テキストや事前の説明だけでは理解が容易ではない事項を効率的に学習させるために、音声ナビゲーションを用いた電子教材を制作することを提案した研究 [3] がある。この研究では、音声ナビゲーションとアニメーション、つまり音声情報と映像情報の2つを用いて通信工学入門科目の電子教材の内容の理解の支援を促すとともに、音声情報の違いによる学習者の反応を測定するために約半分の教材はシミュレーションに関する技術的解説とシミュレーション操作方法の両方を音声でナビゲーションを行った。

また、残りの半分の教材は、シミュレーション操作方法のみを音声でナビゲーションを行った。本研究では、映像情報は用いていないが音声ナビゲーションを用い理解を促すと言う点で共通している。

3 テキスト分析

3.1 テキスト分析のための統合環境 TETDM

レビューやアンケート調査の記述部分、SNSの書き込み、ネット掲示板の書き込みといった様々なテキストを分析するための分析ツールを提供し、これらを統合的に扱えるようにすることで分析に没入し新たな発想を得られる環境である TETDM(Total Environment for Text Data Mining)[4][5] がある。TETDMでは、レビューやアンケート調査の記述部分、SNSの書き込み、ネット掲示板の書き込みといった様々なテキストを分析するための分析ツールを提供し、これらを統合的に扱えるようにすることで分析に没入し新たな発想を得られる環境の構築を目指している。

3.2 テキスト分析の手順と流れの概要

本節では、テキスト分析の手順と流れの概要について述べる。テキスト分析の流れの図を図1に記す。

テキスト分析手順にはデータ分析プロセスと、知識創発プロセスがある。

データ分析プロセスは、データを用意するところからデータから得られた事実を集めるまでの流れを表す。「0. 蓄積データ」で手元に蓄積されているデータから、「1. 分析目的の決定」で分析の目的を決定する。決定した目的に対して、必要なデータが蓄積データだけでは不足している場合、「2. データ収集」で分析に必要なデータを集める。そして、集めたデータを「3. データ整形」で分析ツールに入力できる形式に整形を行い、

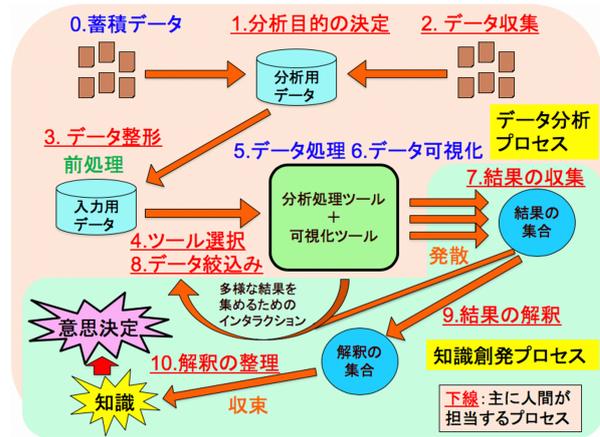


図 1: テキスト分析の流れ

「4. ツール選択」で分析に用いるツールを選択し、「5. データ処理」と「6. データ可視化」で、データ分析処理とその結果の出力の表示を行う。「7. 結果の収集」で分析結果の中で注目する結果を集め、「8. データ絞り込み」は、より多様な結果を集めるために行う。

知識創発プロセスは、データから得られた事実を、意思決定に活用できる知識としてまとめるまでの流れを表す。「7. 結果の収集」で結果を集め、「9. 結果の解釈」で結果から分かる事実を解釈として取り出す。「10. 解釈の整理」で集めた解釈を整理して知識としてまとめる。

4 テキスト分析のための音声ナビゲーション

本章では、音声ナビゲーション機能の概要と前節で示したテキスト分析手順のそれぞれ手順における音声ナビゲーションの実装の方針を述べる。この音声ナビゲーションに従ってテキスト分析を進めていくことでユーザはテキスト分析中につまづきにくくなり、スムーズにテキスト分析を進められると共に分析手順の理解を支援すること目的として音声を準備した。

4.1 音声ナビゲーションの発話条件

本節では、音声ナビゲーションの発話条件について述べる。図2に本システムの構成図を記す。発話条件としては、前章で述べたテキスト分析手順の順番で分析が進んで行くものとして、それぞれの手順に必要なボタン操作やテキストの入力が行われることで、図2のようにその場面に沿った音声ナビゲーションが発話されるようになっている。例えば、マイニングツールをセットした場合には、セットしたツールの使い方の説明や着目すると良いポイントをナビゲーションとし

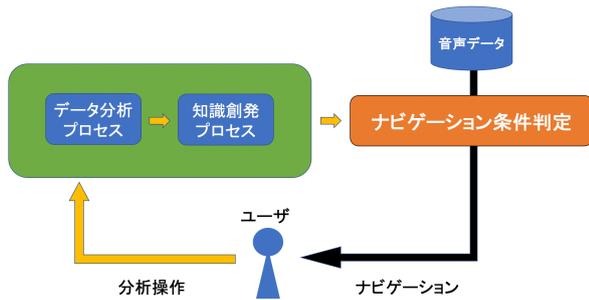


図 2: 音声ナビゲーション機能のシステム構成図

分析手順	つまずきやすいポイント	つまずきを解消するナビゲーション
1.分析目的の決定	分析の目的が決められずテキスト分析を始められない	分析を試しながら目的を決める感じでもいいと思います。
2.データ収集	テキストデータの入力方法がわからない	メニューウィンドウのファイルボタンから、テキストを入力して下さい。
3.データ整形	入力されたテキストが分析に適した整形がされておらず良い結果が得られない	文の数が少ないように思います。テキストデータに句点を追加するか、キーワード設定で、句点の記号を設定して下さい。
4.ツール選択	最初どのツールを使ったら良いかわからない	処理ツール「テキスト評価アプリケーション」の「単語情報」ボタンを押して、ツールをセットしました。
5.データ処理・ 6.データ可視化	選んだツールの使い方、見方がわからない	選んだツールの説明を行う。
7.結果の収集	結果の見つけ方がわからない	結果を見つけるコツは予想と異なる出力に着目することです。
8.データ絞り込み	データ絞り込みを行うタイミングが分からない	データの絞り込みを行っていませんか?
9.結果の解釈	結果の解釈の書き方が分からない	結果の解釈は、結果と解釈の登録ウィンドウで、入力した結果から考えられることを解釈として入力して行きます。
10.解釈の整理	解釈の統合の際どこに着目しながら進めていくかわからない	意味が似ているなど、何らかの共通点をもつ解釈同士を結合させていきます。薄い線で結ばれた解釈には、同じ単語が使われているので参考にして下さい。

図 3: テキスト分析の各手順におけるつまずきやすいポイント

で発話する。つまずきやすいポイントとそれに対する音声ナビゲーションの例を図 3 に記す。

4.2 音声ナビゲーションの内容

本節では、それぞれのテキスト分析手順毎に音声ナビゲーションの内容を述べる。

4.2.1 分析目的の決定

「分析目的の決定」では、分析の目的を決めていないユーザーに対して「分析対象から知りたいことは何か、を考えると決めやすいですよ。また、分析を試しながら目的を決める感じでもいいと思います。」とナビゲーションすることで分析の目的を決めにくい分析初心者に対して目的決定の促進と決定方法の説明を行う。

4.2.2 データ収集

「データ収集」では、テキスト分析に用いるデータが足りない場合に、データを集めてくるように促せるように意図してナビゲーションを行う。例えば、「ではまず、分析に必要なデータを集めてきてね」とナビゲ-

ーションを行うことでデータ収集を促す。また、TETDM のテキストエディタに文章が入力されていない時はテキストの入力方法を説明した上でテキスト入力を促す音声ナビゲーションを行う。

4.2.3 データ整形

「データ整形」では、入力された文章の文の数やセグメントの数の評価を行うことを意図してナビゲーションを行う。評価方法は、文の評価は、文の文字数の平均が 100 文字以上の場合、文ごとの文字数が多すぎるので「テキストデータに句点を追加するか、キーワード設定で、句点の記号を設定してください」とナビゲーションを行うことで文ごとの文字数を減らし、文ごとの特徴をより顕著に表れやすくする。

4.2.4 ツール選択

「ツール選択」では、表 1 の音声ナビゲーションを各発話条件に従って行う。これらの音声は、分析に用いるマイニングツールの選択で初心者が行き詰まらないように意図してナビゲーションを行う。具体的には、「目的に応じたツールを、いろいろと選んで試して下さい」とナビゲーションを行うことでユーザーの分析の目的に沿ったものを簡単にパネルにセットできる。これらのツールセットは、TETDM に実装された処理ツールと可視化ツールを選択した目的に応じて最適な組み合わせで自動でセットを行う。

4.2.5 データ処理・データ可視化

「データ処理・データ可視化」では、これらの音声は、マイニングツールを選ぶことができたが、その使い方がわからず行き詰まってしまうように意図してナビゲーションを行う。具体的には、主題関連語評価が選択された場合「文章中の単語の主題との関連を評価します。ボタン操作で段落ごとに評価することができます」とナビゲーションを行うことでツールの簡単な使い方や見方の説明を行う。

4.2.6 結果の収集

「結果の収集」では、結果の見つけ方がわからず、初心者が行き詰まってしまうように意図してナビゲーションを行う。具体的には、何か気になる結果があった時は、メニューウィンドウの知識創発の結果ボタンを押して、結果と解釈を書き込みましょう。結果を見つけるコツは、予想とは異なる出力に着目することです」とナビゲーションを行うことで、結果の見つけ方を説

表 1: 「ツール選択」における音声ナビゲーションの内容

項番	音声ナビゲーションの内容	音声ナビゲーションの発話条件
4-1	それでは、分析のためのツールをセットしましょう。	ツール初回セット時
4-2	分析では、特徴的な、単語、文またはセグメントに着目します。	4-1 の音声が終わりました
4-3	単語に着目しますか? (Yes or No)	4-2 の音声が終わりました
4-4	処理ツール「テキスト評価アプリケーション」の「単語情報」ボタンを押して、ツールをセットしました。	4-3 の選択肢で「Yes」が選択された
4-5	文やセグメントに着目しますか?	4-3 の選択肢で「No」が選択された
4-6	処理ツール「テキスト評価アプリケーション」の「文・セグメント情報」ボタンを押して、ツールをセットしました。	4-5 の選択肢で「Yes」が選択された
4-7	起動モードをライトモードにして、メニューウィンドウの「目的」ボタンを押すと、目的に応じたツールを選びやすくなります。	4-5 の選択肢で「No」が選択された
4-8	ライトモードに変更しますか? (Yes or No)	4-7 の音声が終わりました
4-9	まずは、処理ツール「テキスト評価アプリケーション」を使って、いろいろなツールを試してみてくださいね。	4-5 の選択肢で「No」が選択された
4-10	どんな分析をしたいですか?	4-5 の選択肢で「No」が選択されたかつモードがライトモード以降
4-11	目的に応じたツールを、いろいろと選んで試してみてください。	4-10 の音声が終わりました

明を行う。また、「結果と解釈の入力だけでなく、その結果を得られたパネル番号と重要度の選択も忘れずに行いましょう」とナビゲーションを行うことで忘れやすい操作を忘れにくくする。

4.2.7 データ絞り込み

「データ絞り込み」では、データ絞り込みを行うタイミングがわからない、または絞り込み機能を忘れている場合に絞り込み機能を使うことを促進する意図でナビゲーションを行う。具体的には、「データの絞り込みを行ってみませんか?データの絞り込みを行って、一部のデータに対する出力を見ることでも、新しい結果を集めることができますよ」とナビゲーションすることで絞り込み機能の存在と使うタイミングを明示する。

4.2.8 結果の解釈

「結果の解釈」では、結果からどのように解釈をすれば良いかがわからなず初心者が行き詰まらないよう意図してナビゲーションを行う。具体的には、「結果の解釈は、結果と解釈の登録ウィンドウで、入力した結果から考えられることを解釈として入力して行います」とナビゲーションを行うことでユーザ自身が感じたことをそのまま記入すればいいと分かるように説明する。

4.2.9 解釈の整理

「解釈の整理」では、結果の解釈を統合させるにあたって、どこに着目して統合させたら良いか分からず行き詰まらないよう意図してナビゲーションを行う。具体的には、「意味が似ているなど、何らかの共通点をもつ解釈同士を結合させていきます。薄い線で結ばれた解釈には、同じ単語が使われているので参考にしてください」とナビゲーションを行うことで統合させるときの着目点を説明する。

5 音声ナビゲーション機能によるテキスト分析支援の検証実験

5.1 実験設定

音声ナビゲーション機能によってテキスト分析をスムーズに行えるかどうかを評価するために、テキスト分析に不慣れな理系の大学生 18 名に、音声ナビゲーション機能を利用しながらレビューテキストの分析を行ってもらった。

まず、TETDM のチュートリアルを一通り被験者に行ってもらい、TETDM の基本的な使い方や最低限のツールの使い方、結果と解釈の書き方を理解してもらった。このチュートリアルでは、あらかじめ用意されたテキストデータを使って、指示された通りに TETDM の操作を行っていくことで、テキスト分析の手順それぞれの基礎的な知識を身につけることができる。

チュートリアル終了後、チュートリアルに要した時間が均等に分かれるように被験者をグループ A、B の 2 グループに分ける。その後 3 つのレビューデータ [6] でテキスト分析を行ってもらった。グループの振り分けとテキストと音声ナビゲーションの使用の有無の振り分けを表 2 に記す。

グループ A、B 共に練習として改善案のアイデアの提案を分析の目的としてワイヤレスイヤホンのレビューデータのテキスト分析を音声ナビゲーション機能なしで行ってもらった。この際に、集めてもらう結果と解釈の個数は 10 個以上、実験時間は 30 分から 1 時間を目安に行った。

2 つ目はスマートスピーカーのレビューデータのテキスト分析を行ってもらった。この際に音声ナビゲーション機能の有効性を調べるためにグループ A は音声ナビゲーション機能なしで分析を行い、グループ B は音声ナビゲーション機能ありで行った。

3 つ目はタブレット端末のレビューデータのテキスト分析を行ってもらった。今回は、グループ A は音声

ナビゲーション機能ありで分析を行い、グループ B は音声ナビゲーション機能なしで分析を行った。

表 2: 実験のグループ振り分け

グループ	手順 1 イヤホン (練習)	手順 2 スマート スピーカー	手順 3 タブレット 端末
A	ナビなし	ナビなし	ナビあり
B	ナビなし	ナビあり	ナビなし

5.2 音声ナビゲーション機能によるテキスト分析のスムーズさの評価

TETDM を操作している時間を除いたテキスト分析全体にかかったグループ毎の平均時間を図 4 に示す。図 4 から、練習テキストでの分析時間平均は、グループ A では 55 分、グループ B では 62 分という結果になった。また、実験時間はどちらのグループでも同様にナビゲーション機能ありで実験を行ったところ 35 分、ナビゲーション機能なしの場合が 42 分という結果が得られた。

この結果から、ナビゲーション機能があるの方が、なしの場合よりも平均 7 分、分析を早く終えることができていることがわかるので、この結果から音声ナビゲーション機能は、入力されるテキストの種類によらずスムーズなテキスト分析を支援できていることがわかった。

一方でグループ B のナビゲーション機能なしの時間がナビゲーション機能ありの後にも関わらずグループ A のナビゲーション機能ありの場合と同じ時間になった。しかし、練習テキストの時間をグループ A, B で比べると B の方が時間が長くテキスト分析が苦手な人が多いと考えられるので、グループ B でも音声ナビゲーションの効果によりテキスト分析の手順理解に役立ったと考えられる。

5.3 音声ナビゲーション機能の効果

分析をスムーズに行えることがわかったが、実際にテキスト分析のどこの部分でナビゲーション機能の効果が大きかったのかを調べるために、一番実験時間に影響が大きいと考えられるツール選択のプロセスに着目して、結果と解釈の登録に関係のないマイニングツールが使われた時間を調べ、音声ナビゲーション機能がある場合となしの場合で分けて結果を昇順に並べたものを図 5 に示す。図 5 の真ん中より左側の結果登録に関係のないツールが使われた時間が少ない場合では、ナビゲーション機能がなしの場合、0 が多い結果が得られた。

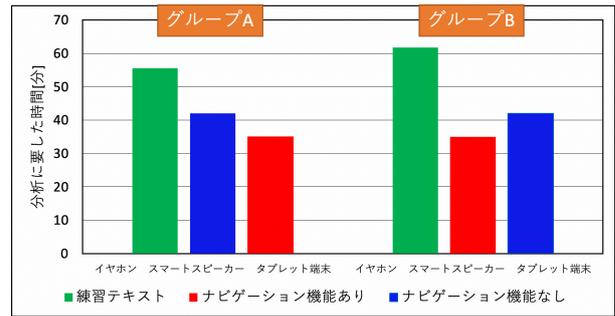


図 4: グループ毎のテキスト分析にかかった平均時間

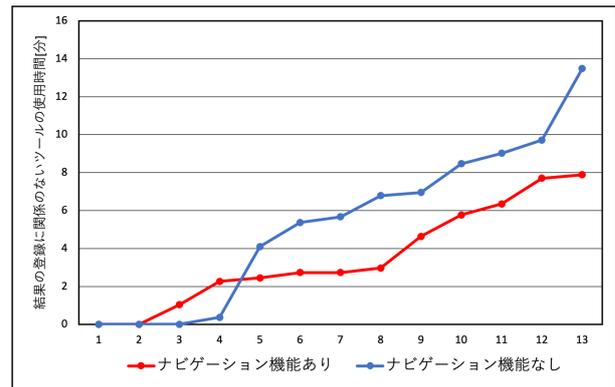


図 5: 結果登録に関係のないツールが使われた時間 (時間順にソート)

これは、分析操作に慣れていて、あまりツールを使わないためナビゲーションの効果小さかった体と考えられる。一方で図 5 の真ん中より右側では、結果登録に関係のないツールが使われた回数は、ナビゲーション機能がある場合がなしの場合よりも少ない結果が得られた。主に結果登録に関係のないツールを、多く使うのは、テキスト分析に慣れていない初心者だと考えられるので、本研究の音声ナビゲーション機能は、テキスト分析初心者に対して特に効果的であることがわかった。また、ツールを多く使うユーザに効果が大きく見られたので、音声ナビゲーション機能で特に効果的であったテキスト分析プロセスは、ツール選択であることがわかった。

6 おわりに

テキスト分析において、スムーズに分析を行えるように音声ナビゲーションを用いて分析の支援を行った。本研究のナビゲーション機能が、テキスト分析を行う際に分析をスムーズに行えるかを検証するために、音声ナビゲーション機能を用いて実際にテキスト分析を行い検証実験を行った。実験の結果から、音声ナビゲーション機能を用いた場合、なしの時に比べて分析にかかる時間が短くなる傾向が得られ、スムーズにテキス

ト分析を行うことができた。また、今回実装した音声ナビゲーションの中で最も効果が大きかったテキスト分析プロセスはツール選択であることがわかった。

今後の課題として、テキスト分析初心者だけでなく基本的な知識を持っているユーザに対しても有益なナビゲーションの実装を検討していきたい。

参考文献

- [1] 保田 明夫: テキスト・マイニングの概要, 電気学会論文誌 C (電気・情報・システム部門誌), Vol.125, No. 5, pp. 682-689(2005)
- [2] 西原 陽子, 中垣内 李菜, 川本 佳代, 砂山 渡: TETDM を用いたテキストマイニング のスキル獲得を支援するためのチュートリアルシステムの開発, 知能と情報, Vol.27, No. 5, pp.771-783(2015)
- [3] 今澤 達哉, 稲垣 貴俊, 野須 潔: 音声ナビゲーション付きシミュレーション型通信工学入門電子教材の制作とユーザビリティ評価, 映像情報メディア学会誌, Vol.59, No.1, pp. 65-68 (2005)
- [4] Total Environment for Text Data Mining (テキストデータマイニング のための統合環境), (URL)<https://tetdm.jp>
- [5] 砂山渡, 高間康史, 徳永秀和, 串間宗夫, 西村和則, 松下光範, 北村侑也: 統合環境 TETDM を用いた社会実践, 人工知能学会論文誌, Vol.32, No.1, NFC-A, pp.1-12 (2017)
- [6] Amazon.com, Inc.,(URL)www.amazon.co.jp

BERT を用いた単語補完による褒め台詞生成システム

Complimentary Text Generation System by Word Completion using BERT

堀 紘之輔 砂山 渡*
Hironosuke Hori Wataru Sunayama

滋賀県立大学 工学部
School of Engineering, The University of Shiga Prefecture

Abstract: Nowadays, many Japanese people lead stressful lives. We think that being praised can release our stress and we have better life. Therefore, this study will picking out a word that is related to a input complimentary word using BERT. Also we prerare often used word group and sometime used word group, purpose of this study is generating various complimentary text.

1 はじめに

近年日本はストレス社会と言われており、日々の生活の中で7割近くの人がストレスを感じているというアンケート結果もある [1]。そんな世の中で褒めてもらうことは日々の活力やモチベーションの向上やストレスの軽減により繋がるのではないかと考えた。しかし、近年では一人暮らしの割合が増えている [2]。そのため身近に褒めてくれる関係の人がいない人も増えてきていると考えられる。そこで単語に対して褒め台詞を出力することができれば、会話ボットのようなものを用いて一人でも自分の行いに対して褒めてもらうことができ、より良い生活を送れるのではないかと考えた。

そこで、本研究では近年発展してきた BERT という学習済みモデルに注目した。学習済みの BERT モデルを用いることで入力とする単語に対してより関連性の高い単語を抽出し、その中から褒められると考えられる単語を抽出することで、入力単語に対する褒め単語を得て、また汎用的でない褒め単語群と汎用的な褒め単語群を用意することで、よりの確かつバリエーション豊かに褒め単語を出力し、それをテンプレートに当てはめることでより相手に嬉しいと思ってもらえる褒め台詞の生成を目的としている。

2 関連研究

2.1 文章生成についての関連研究

機械学習による文章生成 [3] がある。この研究では、word2vec と LSTM を用いることで、228965 語の語彙を持つ文章生成システムを構成した。ユーザから入力された文章を MeCab による形態素解析、分かち書きを行い、得られた形態素を word2vec に入力して文章の分散表現を獲得する。対話文の学習をさせた LSTM モデルに、その分散表現を入力することで返答文を生成する。結果として、この文章生成システムで雑談が可能であることが示せた。

また GPT-2 の転移学習によるキーワードを考慮した広告文生成 [4] がある。この研究では、事前学習を行うことでデータに含まれるキーワードを抽出し、キーワードを条件づけて広告文を生成する学習を行なっている。また少ないデータで優れた性能を示すため、日本語データで事前学習を行った GPT-2 を用いて、広告データで転移学習を行っている。

他には批判的思考を促進する対話文生成に関する研究 [5] がある。この研究では、入力された文章から属性語を抜きだし、属性語同志の共起する確率と属性語と評判語の共起する確率を用いて関係の深い属性語を導く。ここで BERT の Next Sentence Prediction を用いて会話の流れを意識した文章を生成する。

そこで本研究では学習モデルに BERT [6] を用いることで単語を補完し、台詞を生成することを試みた。

*連絡先： 滋賀県立大学工学部 砂山 渡
〒 522-8533 滋賀県彦根市八坂町 2500
E-mail: sunayama.w@e.usp.ac.jp

表 1: 褒め単語の一覧

汎用的でない褒め言葉	積極的, 一流, ナイス, 大人, プロ, 優秀, 優雅, 華麗
	達人, 多才, 天才, 美味, 個性的, 美人, 繊細, 豊富
	器用, 紳士, 親切, 誠実, 男前, 献身的, 高価, 裕福
	ユニーク, 嬉しい, 新鮮, 熱心, 正確, 頑丈, 丈夫, 幸せ
	カッコいい, スマート, 優しい, 明るい, 面白い, うまい
	羨ましい, 美しい, 強い, おいしい, 美味しい, 美味い
汎用的な褒め言葉	鋭い, 重要, 強力, 早い, 速い, はやい, 綺麗, きれい
	凄い, すごい, 素晴らしい, すばらしい, 立派, 最高
	うまい, 上手い, 上手, いい, 良い, 最高, 完璧, 立派
	素敵, 流石, さすが

2.2 本研究における「褒める」の定義

広辞苑 [7] によると、「褒める」は、祝う。ことほぐ。祝福する。(同等または目下の者の) 行いを評価し、よしとしてその気持を表すとある。また、褒めの役割や機能については、公的評価を行う評価行為としての機能や社会関係の創造あるいは保持という社会的潤滑油としての役割がある [8]。そこで本研究では、相手を喜ばせる言葉を褒め言葉として、相手を喜ばせる単語を褒め単語とする。

本研究の褒め単語として、後述する提案システムに 100 の名詞を入力としたときにそれぞれの上位 20 位以内に出力された単語のうち、著者の一人が褒め言葉になりうると考えた単語を褒め単語として用いる。また、それらの集めた褒め単語について、100 の名詞の半数以上において文脈に違和感なく用いられる褒め単語を「汎用的な褒め単語」、その他の褒め単語を「汎用的でない褒め単語」とする。本システムで用いた褒め単語一覧を表 1 に示す。

3 BERT を用いた単語補完による褒め台詞生成システム

本章では、褒め台詞生成システムについて、システムの概要、作成した機能について述べる。

3.1 システムの概要

BERT を用いた単語補完による褒め台詞生成システムの構成図を図 1 に示す。このシステムでは、まず自分に関する褒めて欲しい「1 単語の名詞」を入力とする。その入力された単語に関する褒め単語を入力用のテンプレートに入れて入力用の文章を作り、BERT にその文を入力することで褒め単語を出力する。あらかじめ用意した出力用のテンプレートに出力された褒め単語を入れて褒め台詞として出力する。

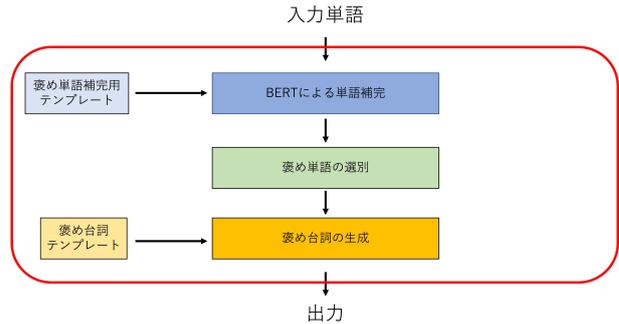


図 1: BERT を用いた単語補完による褒め台詞生成システムの構成図

3.2 BERT を用いた単語補完による褒め台詞生成システムへの入力

本システムの入力は、褒めてほしい「1 単語の名詞」とする。「映画鑑賞」(「映画」+「鑑賞」)のような複合名詞は対象としていない。これは BERT では文章を 1 単語ごとに分割して学習しているため、2 単語以上の文をそのまま入力としてもその文を 1 単語のように扱ってしまい、またそのような単語は学習されていないため、適切な出力がされないためである。

3.3 BERT による単語補完

3.3.1 BERT への入力用のテンプレート

表 2 に各テンプレートに対して 100 単語の入力をした際の出力がされた入力の単語数と、出力された褒め単語の種類と、出力された褒め単語の平均順位をまとめたものを示す。ただしテンプレート内の〇〇は入力する単語であり、**の部分の単語を補完して褒め単語を出力する。BERT への入力用のテンプレートの作成にあたって以下の点に注目した。

- より多くの単語を出力されるか
- より様々な褒め単語が出力されるか
- より上位で褒め単語を出力することができるか

これらの結果から「〇〇はとても**です。」と「あなたの〇〇はとても**です。」の 2 つのテンプレートが適していると考えられ、また「あなたの〇〇はとても**です。」の方がより入力単語に対して適切な褒め単語が出力されていたと考えたため、本システムでの BERT への入力用のテンプレートは「あなたの〇〇はとても**です。」とした。

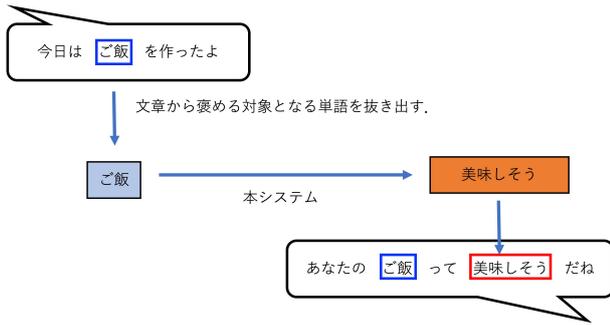


図 4: チャットボットへの応用例

して出力することで、自然な会話の流れで相手を自動で褒めることが可能であると考えられる。また音声で読み上げする機能をつけることで人に褒められるということに近づき、より嬉しく感じるのではないかと考える。

4 BERTを用いた単語補完による褒め台詞生成システムの評価実験

本章では、実装した良好な人間関係の構築に向けた文章作成支援機能が実際に良好な人間関係の構築に有用であるかを検証するために行った実験について述べる。本実験では、作成したBERTを用いた単語補完による褒め台詞生成システムにより得られた褒め台詞が褒め台詞として成り立っているか、またより嬉しい褒め台詞の生成ができるかの評価を目的とする。

4.1 実験手順

本実験は、BERTを用いた単語補完による褒め台詞生成システムと word2vec の類似度を用いて得た褒め台詞を比較する実験を行った。具体的には、被験者には自分に関する褒めてほしい1単語の名詞を入力してもらい、作成したBERTを用いた単語補完による褒め台詞生成システムにより得られた褒め台詞と、東北大学の公開している word2vec の学習済みモデル [10] を用いて、褒める対象の単語と褒め単語との類似度から得られた褒め台詞をそれぞれ評価してもらい、作成したシステムがより自然で相手を喜ばせる褒め台詞を出力できるかを評価する実験を行った。実験手順を以下に述べる。

1. 自分に関する褒めてほしい「1単語の名詞」を入力する。
2. BERTを用いた単語補完による褒め台詞生成システムで生成された褒め台詞と word2vec の類似

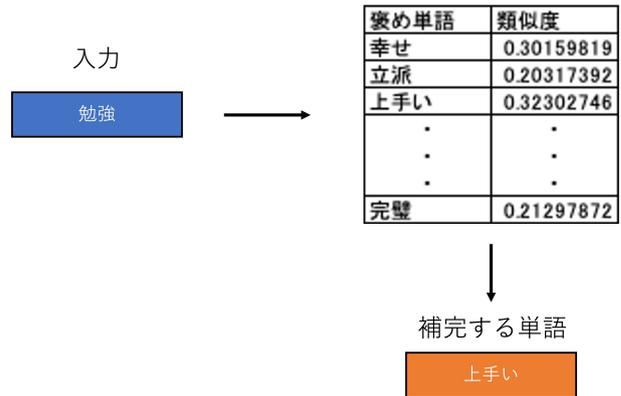


図 5: 比較システムの出力例

度を用いて得た褒め台詞をそれぞれ「嬉しい」、「嬉しくない」、「日本語が変」の3つうちのいずれかで回答する。褒め台詞の出力がされなければ手順1に戻る。

3. どちらの褒め台詞にも「嬉しい」を選んだ場合には、どちらの方が「より嬉しい」かを選択する。
4. 実験手順1~3を20単語について行う。「褒められません。」と出力される単語が入力された場合は、別の単語を再度入力する)

比較システムの出力例を図5に示す。比較するシステムでは、入力された単語と褒め単語の類似度を網羅的に算出し、その中で最も類似度が高いものを補完する単語として本システムと同様に「あなたの○○って**ね。」のテンプレートに入れ、褒め台詞を生成する。

4.2 実験結果と考察

「嬉しい」と回答した褒め台詞の数を図6に示す。この図から提案システムの方が「嬉しい」と回答した数が多くなった被験者の数が8人中7人となり、より多くの単語に関して褒める台詞の生成ができたと考えられる。比較システムの方が「嬉しい」と回答した数が多くなった被験者に関しては、比較システムで「嬉しい」と回答した数が他の被験者と比べて一番多くなっており、また提案システムで「嬉しい」と回答した数は他の被験者と大きな差はなかったため、この被験者の入力として多かった趣味に関する入力では比較システムが優秀な結果を示すことができるために、提案システムより比較システムの方が「嬉しい」の数が多くなったと考えられる。

「より嬉しい」と回答した褒め台詞の数(ただし、片方のシステムのみ「嬉しい」を選択した場合を含む)を

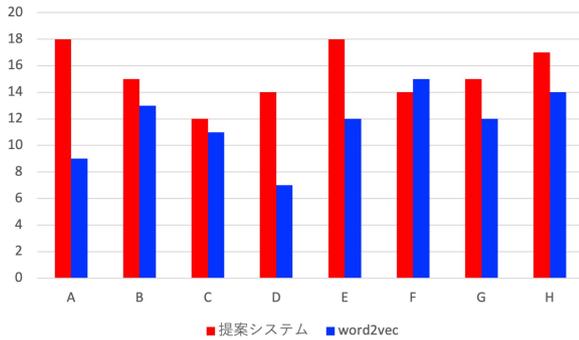


図 6: 「嬉しい」と回答した褒め台詞の数

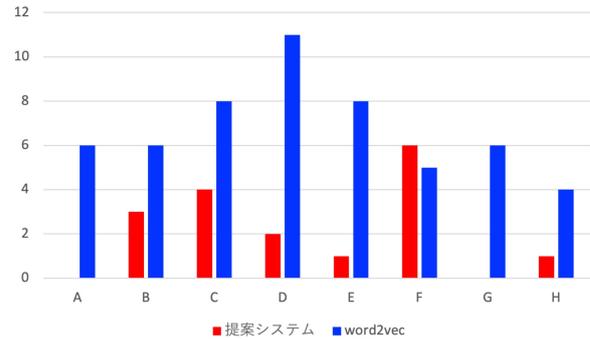


図 8: 「日本語が変」と回答した褒め台詞の数

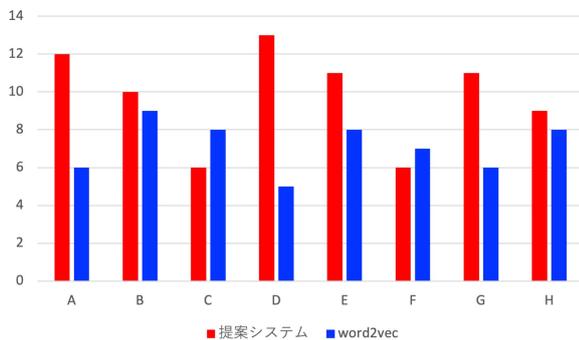


図 7: 「より嬉しい」と回答した褒め台詞の数

図 7 に示す。この図から提案システムの方が「より嬉しい」と回答した数が多かった被験者が 8 人中 6 人となり、より嬉しい褒め台詞を作ることができている。比較システムの方が「より嬉しい」の数が多かった被験者に関して、入力単語が容姿に関するものが多くっており、そのため出力される単語が「きれい」、「美しい」が多くなり褒め台詞が単調となってしまったことが原因と考えられる。

「日本語が変」と回答した褒め台詞の数を図 8 に示す。この図から提案システムの方が「日本語が変」と回答した数が少なくなった被験者の数が 8 人中 7 人となり、また「日本語が変」と回答した数が少なくなった被験者の数が 2 人いたことがわかる。これにより提案システムではより違和感の少ない褒め台詞を生成することができたと考えられる。提案システムの方が「日本語が変」と回答した数が多かった被験者に関して、この被験者の入力は趣味に関するものが多く、趣味に関しては BERT 補完用のテンプレート「あなたの〇〇はとても**です。」が適切でないと考えられた。そのため、BERT 補完用のテンプレートも複数用意し、入力に対して使い分けを行ったり、それぞれのテンプレートで生成した褒め台詞を BERT などを用いて評価して、

表 3: 各被験者の評価

被験者	提案システム			比較システム		
	より嬉しい	嬉しい	日本語が変	より嬉しい	嬉しい	日本語が変
A	12	18	0	6	9	6
B	10	15	3	9	13	6
C	6	12	4	8	11	8
D	13	14	2	5	7	11
E	11	18	1	8	12	8
F	6	14	6	7	15	5
G	11	15	0	6	12	6
H	9	17	1	8	14	4
平均	9.75	15.38	2.13	7.13	11.63	6.75

最も文章として適当なものを最終的な出力とするなどで解決できると考えられる。

各被験者の評価を表 3 に示す。この表から平均は全ての数字に関して提案システムの方が優れていることがわかる。特に「日本語が変」の数は提案システムの方が比較システムに対して 3 分の 1 以下の数となり、特により違和感の少ない褒め台詞を生成することが得意であると考えられる。比較システムで「日本語が変」の数が多かった原因として、比較システムではあくまで 2 単語の類似度から褒め単語を補完しているため、文脈に関係なく類似度の高い褒め単語が補完されるためであると考えられる。例えば、「高校」と「プロ」の 2 単語はスポーツの分野において密接な繋がりがあると考えられるが、「あなたの高校ってプロね。」は日本語が変であり、文脈の理解ができていない。そのため、逆に提案システムでは BERT を用いることで文脈を理解して褒め台詞を補完することができていると考えられる。

以上のことから、作成した BERT を用いた単語補完による褒め台詞生成システムは、より自然な褒め台詞を生成することが可能となったと考え、本システムの目的を達成できた結果となったと考える。

5 おわりに

BERT を用いた単語補完による褒め台詞生成システムを実装した。また、BERT を用いた単語補完による褒め台詞生成と word2vec を用いて、褒める対象の単語と褒め単語との類似度から得られた褒め台詞をそれぞれ評価、比較してもらう実験を行った。

この実験において、作成した BERT を用いた単語補完による褒め台詞生成システムから得られた褒め台詞は word2vec の 2 単語間の類似度から得られた褒め台詞よりも、「嬉しい」と思わせる台詞を生成できる可能性が高いという結果となった。課題としては、出力する褒め台詞のバリエーションを増やす必要がある。そのため登録する褒め単語の追加や、汎用的でない褒め単語として登録されているもの中でも出力される確率が高い単語の汎用的な褒め単語への変更、最終的な単語の出力の調整などの検討が必要である。

提案する褒め台詞の生成システムを用いることで、今後より良い暮らしへの支援が可能となることに期待したい。

参考文献

- [1] マイボイスコム：ストレスに関するアンケート調査(第7回), <https://myel.myvoice.jp>(2022年1月28日確認)
- [2] 総務省統計局：平成27年国勢調査, 我が国人口・世帯の概観, 第10章世帯数, 世帯の家族類型, p.121 (2015)
- [3] 深井宏剛, 松澤智史：機械学習による文章生成, 情報処理学会第82回全国大会, 1T-02, pp.2-415-2-416 (2020)
- [4] 大曾根宏幸, 張培楠：GPT-2の転移学習によるキーワードを考慮した広告文生成, 第35回人工知能学会全国大会, 2D4-OS-7b-03 (2021)
- [5] 吉田聖, 馬強：批判的思考を促進する対話文生成に関する研究, DEIM2020, A6-3 (2020)
- [6] Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, In proceedings of the 2019 Conference of North America Chapter of Association for Computational Linguistics: Human Language Technologies, Vol.1, pp.4171-4186 (2019)
- [7] 新村出：広辞苑第七版, 岩波出版 (2018)
- [8] 澤口右京, 渋谷昌三：「ほめ」に関する心理的研究の動向, 目白大学心理学研究, No.10, pp.93 - 104 (2014)
- [9] 東北大学乾研究室：Pretrained Japanese BERT models released / 日本語 BERT モデル公開, <https://github.com/cl-tohoku/bert-japanese> (2022年1月27日確認)
- [10] 東北大学乾研究室：日本語 Wikipedia エンティティベクトル, <https://github.com/single-tongue/WikiEntVec>(2022年1月27日確認)

複数 BERT モデルで作るツールによって コメントの可読性向上を試みる

Using multiple BERT models for Improving the Readability of YouTube Comments

伊熊秀磨 ダビド・ラマモンジソア

Hidemaru Ikuma, David Ramamonjisoa

岩手県立大学ソフトウェア情報学部

faculty of Software and Information Science Iwate Prefectural University

Abstract: Nowadays, it is common to read the comments along with the video when watching a YouTube video. While reading the comments can provide new insights and empathy, some comments can be insulting or offensive. YouTube has its own way of displaying comments based on the top ratings or newest. In this study, we created a comment analysis tool that adds five analysis criteria to comments: harmful, spam, sentiment, sarcasm, and constructive. In this paper, we propose a tool and a framework to analyze and visualize comments on YouTube. Finally, we present some evaluation results.

1. はじめに

ユーザのコメントには、そのコンテンツに対する貴重な情報が含まれており、様々な利用方法が提案されている。また、それらの表示方法についても、ランキング化やフィルタリング、要約など様々な研究がされている[1]。本論では現在の YouTube のコメント表示方法に疑問を持ち、より良い改善方法があるのではないかと考えた。ユーザによる評価では、有害なコメントを判断することができず、多様なコメントを素早く理解することも難しい。そこで、BERT (Bidirectional Encoder Representations from Transformers) を用いた学習を行うことで、コメントのフィルタリングとソートを行い、より見やすいコメント欄を作ることを目的とした。

2. 手法

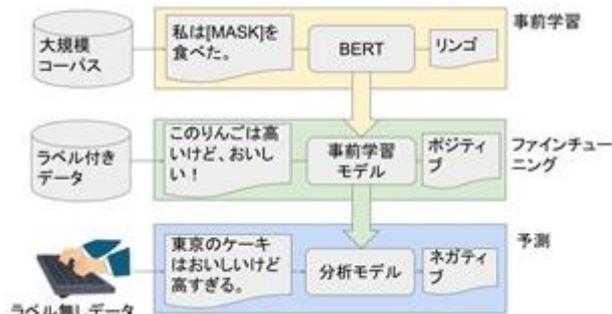


図 1 モデル作成の流れ

コメント分析には、Google が開発した BERT-Tiny 事前学習モデル[2]を使用した。有害[3]、スパム[4]、センチメント[5]、皮肉[6]、建設的[7]の 5 つのデータセットを用いて、図 1 のようにそれぞれファインチューニングを行ったモデルを分類モデルとしている。

ツールの仕組みは図 2 のようになっている。最初に、ツールで動画 URL の入力を行うと YouTube DataAPI からコメント情報を取得し、データベースに保存する。次に、分析モデルはデータベースのコメントを分析し、結果をデータベースに返す。最後に、ツールはデータベースの情報を読み込み表示する。ユーザは 5 つの分析基準を用いてフィルタとソートをされたコメントを読むことができる。



図 2 ツールの仕組み

3. データについて

有害[3]、スパム[4]、センチメント[5]、皮肉[6]、建

設的[7]の5つのデータセットは表1のように収集されたラベル付きコメントである。

表1 コメントデータセット

	Web サイト	件数
有害	Wikipedia	223,549
スパム	YouTube	1,956
センチメント	Twitter	1,600,000
皮肉	TheOnion、 HuffPost	28,619
建設的	SFU News	12,000

4. ファインチューニングの設定

コメント分析には BERT モデルを用いている。BERT モデルとは、Bidirectional Transformer encoder を Layer としたモデルである。事前学習には Masked Language model と、Next Sentence prediction が行われている。

BERT モデルは keras ライブラリを用いて次のように設定した。テキストデータの埋め込みに事前学習モデルの BERT-Tiny モデル[2]を使用した。隠れ層 128 ユニット、Layer2 層、Attention head2 のモデルである。SpatialDropout1D で 10%の単語を無視した後、それらを Dense にて sigmoid 関数で全結合している。loss 関数には binary_crossentropy、optimizer には adam、batch_size は 512 としている。

4.1 学習回数の模索

データセットを学習データ 9 割、テストデータ 1 割に分割し、最も良いスコアとなる学習回数を模索した。また模索の際、最後の結合時に、先頭 1 ユニットを用いる lunit スタイル (lunit) と、全てのユニットを用いる sequence スタイル (SEQ) の 2 パターンを用意して模索した。有害、スパム、皮肉、建設的データセットは 100epoch まで行い、センチメントデータセットは 10epoch まで行った。

4.2 学習回数の決定

模索で得られた正答率のスコアを参考に、各モデルの学習回数と結合方法を表2のように決定した。

表2 学習回数と結合方法

	有害	スパム	センチメント	皮肉	建設的
学習	1 回	10 回	5 回	2 回	2 回
結合	lunit	SEQ	SEQ	SEQ	SEQ

それぞれのデータセットのモデルを評価した結果は表3のようになっている。適合率、再現率、F1 値は 2 クラスのマクロ平均を算出している。センチメントのスコアが他に比べて低い。これは、データセットが大規模であり、十分な学習を模索することが

できなかったことが原因である。

表3 各モデルのスコア

	有害	スパム	センチメント	皮肉	建設的
正答率	0.956	0.954	0.833	0.897	0.927
適合率	0.908	0.954	0.833	0.897	0.926
再現率	0.836	0.953	0.833	0.896	0.929
F1 値	0.868	0.953	0.833	0.896	0.927

5. 結果と考察

5.1 完成したツールの説明

5つのモデルを使用することで、図3のWebサイトが完成した。左から順番に投稿日時、投稿者、コメント、グットボタンの数、有害、センチメント、皮肉、スパム、建設的となっている。このうち、最後の5つはモデルによって予測した値を表示している。

日時	投稿者	コメント	有害	スパム	センチメント	皮肉	建設的
2021-11-24 22:53:05	[top]	[discuss]UPDATE: We managed to get Blender to run on the CMP card after posting this video. The results are not very good due to limited memory and PCIe bandwidth. Here are render time results for the BMW blender demo. Lower is better.	31.52	0%	15%	99%	100%
2021-11-25 16:42:34	[top]	[discuss]I personally don't believe you mining is the future and can be damaging to the climate with its power consumption. Cpu mining is more eco friendly and some really good projects popping up. Been watching verus coin as it appears to be eco friendly and one of the more advanced cryptos out there and fixing a lot of problems around crypto. Lead	1	0%	38%	98%	100%
2021-11-25 15:11:27	[top]	[discuss]@K04th No sorry these and the A1000 are mainly meant to be used in a VM. To be able to use the entire cones on one VM you have buy a special license from Nvidia for V GPU. If not you can only use pass thru on hyper V and lower end VMware licenses you can only pas thru 4 core clusters. If you upgrade to full VMware Enterprise you can pass	2	0%	15%	2%	100%
2021-12-14 03:27:30	[top]	[top]If you smart guys want to explain how a graphics card can "mine" a "currency?" Because it's all a scam in my opinion. There's no physical way to obtain a Bitcoin. You just believe in this weird fake currency. I can't bring myself to invest in fake money. Lol I'll buy	0	0%	22%	10%	100%
2021-12-13 14:19:40	[top]	[top]So basically, its not a profitable mine. Until you resell it. So whos buying these used? I would love to know just how much this brings in. Combine that with the difficulty increase I wonder how long it will be profitable for. (once you reach the break even point. A great test would be to run this til its no longer making anything. sell it and see where you end up.	0	0%	27%	0%	100%
		[top]SO DISGUSTING THAT THINGS LIKE THIS EXIST. STEAL FROM THE RICH!!! IF YOU					

図3 分析ツールの表示画面

discussion のボタンを押すことで、図4のように対話の様子を表示することができる。

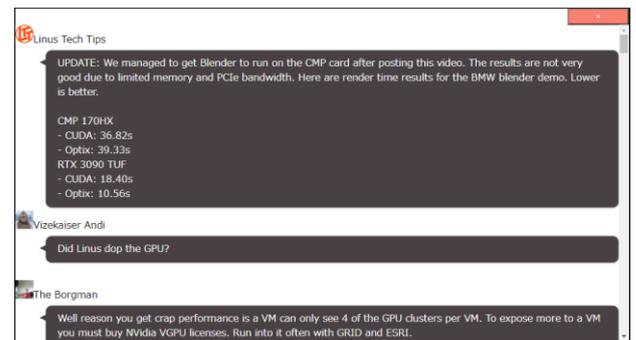


図4 discussion の表示

図5のように、ツールはソートや5つの予測値に対してフィルタを行うことができ、その分布やトップレベルのコメントとリプライのコメントの割合も表示する。フィルタは予測値の下限と上限を設定することで、その範囲内のコメントのみを表示する。分布は x 軸と y 軸に 5 つの分析基準を指定することで 2 次元座標にプロットされる。プロット点にマウスオーバーすることでコメントの一部が表示され、クリックすることで詳細の場所へジャンプする。

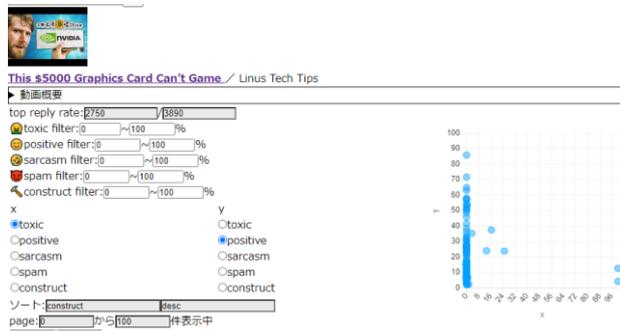


図 5 分析ツールの操作画面

5.2 ツールの使用感

ツールを本研究室の3年生に使用してもらった結果、以下のようなフィードバックを得た。

- コメント件数の多いコメントを取得した場合、コメントの表示が途中で止まってしまう。
これは YouTubeAPI に起因する問題であるため、API の取得上限解放を申請することで解決する。ただし、それでも取得が途中で止まってしまった場合、何らかの代替手段が必要である。
- グットボタンが多い順にできるか
できるが、この質問は操作性が分かりにくいという問題に起因している。より直感的な操作方法の検討や操作マニュアルの表示が必要である。
- ランダムにコメントを取得したい
ソートやフィルタの機能は実装していたが、ランダムの検討はしていなかった。
- 「F*CK」が入っていると必ず toxic が 99% になる

F*CK は肯定的であれ否定的であれ少し汚い表現であるため、ツールが良い働きをしている。

また、spam の予測に関してはデータセットが少ないということもあり、spam 以外のコメントも spam であると判定してしまう傾向にある。

6. おわりに

本稿では、複数 BERT モデルで作るツールによってコメントの可読性向上を試みた。YouTube のコメント表示方法と異なる手法を用いることで、新しい視点で新しい知識やコミュニティと出会うことができる。しかし、表 3 のスコアで示した通り予測モデルの精度が悪いため、フィルタとしての役割は果たせていない。より良いモデル構築とデータセットの収集が必要である。

参考文献

- [1] Aggregating and Analyzing Articles and Comments on a News website <http://p-www.iwate-pu.ac.jp/~david/>

[publications/DavidRama-28June2016-LNAI-Final.pdf](https://arxiv.org/abs/1606.04934)

- [2] google-research/bert: TensorFlow code and pre-trained models for BERT <https://github.com/google-research/bert>
- [3] Toxic Comment Classification Challenge | Kaggle <https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge>
- [4] YouTube Spam Collection Data Set | Kaggle <https://www.kaggle.com/lakshmi25npathi/images>
- [5] Sentiment140 dataset with 1.6 million tweets | Kaggle <https://www.kaggle.com/kazanova/sentiment140>
- [6] News Headlines Dataset For Sarcasm Detection | Kaggle <https://www.kaggle.com/rmisra/news-headlines-dataset-for-sarcasm-detection>
- [7] Constructive Comments Corpus (C3) | Kaggle <https://www.kaggle.com/mtaboada/c3-constructive-comments-corpus>

位置情報に基づいた画像分類モデルの選定による 看板画像の分類精度向上

Improving Classification Accuracy of Signboard Images by Selecting Location-Based Image Classification Models

中道 翔大* 松下 光範
Shota Nakamichi Mitsunori Matsushita

関西大学総合情報学部
Faculty of Informatics, Kansai University

Abstract: 本研究の目的は、看板から店舗の詳細情報を取得するシステムの実現である。これまでに、携帯端末のカメラで撮影した看板画像から店舗を判別し、ARを用いてその店舗の詳細情報を表示する手法が提案されている。しかし、このシステムは識別可能な店舗数が限られており、対象店舗数を拡大する場合、分類する店舗数の増加に伴って識別精度が低下する懸念がある。そこで本稿では、対象店舗の所在地をもとに予め複数の画像分類モデルを構築し、ユーザの位置情報に基づいて使用するモデルを選定して識別を行うことで、分類精度を向上させる手法を提案する。

1 はじめに

看板とは、「屋外広告物法」により定められた条件を満たし、商店などが店名・業種・商品名などの情報を道路利用者の目に付きやすいように屋外に掲げたものであり、各店舗は街中を歩く人々に対して来店を促すために看板を掲示する [10]。人々は、目に入った看板に描かれている内容から店舗の種類を判別し、どの店に入るかななどの意思決定を行っている。加えて、スマートフォンなどの携帯端末を用いた、看板に記述されている店舗名をもとに行うテキスト入力による Web 検索や、位置情報をもとに行う付近の店舗検索によって、その店舗の商品や口コミなどの詳細情報を取得できる。

看板の多くはその店舗が位置する地域の言語で記述されている。そのため、看板の文字がユーザの母国語でない場合、言語障壁によって看板から店舗の情報を得ることが困難となる。草書体のような読みづらい書体で記述されている看板の場合でも、同様のことが言える。テキスト入力による Web 検索を行う場合でも、ユーザが文字を自力で入力できることが前提となっているため、看板の文字が理解できない場合、店舗の詳細情報の取得が困難となる。位置情報を用いた検索によって周辺にある

自分の条件に合った店舗が見つかったとしても、看板の文字が読めない場合、多くの店舗が密集している中から目的の店舗を発見するには手間がかかる。そのため、人々が目的の店舗を探す上で、その店舗の看板の文字が理解可能であるかは重要な要素であると言える。

本研究では、ユーザが周囲の文字が理解できない状況にいる場合でも、看板から目の前の店舗の情報を容易に取得できるシステムの実現を目的としている。北村らは、携帯端末のカメラで撮影した看板画像から、機械学習を用いた物体検出及び画像分類によって店舗を判別し、ARを用いて店舗の詳細情報を重畳表示するシステムを開発している [8]。しかし、北村らのシステムでは特定の地域に位置する 15 店舗のみを対象としているため、適用範囲が限られていた。このシステムの対象を拡大する場合、分類する店舗数の増加に伴って精度が低下する懸念がある。この問題を解決するため本稿では、対象店舗の所在地をもとに予め複数の画像分類モデルを構築し、ユーザの位置情報に基づいて使用するモデルを選定することで、分類精度を向上させる手法を提案する。

2 関連研究

本章では、看板に記述されている文字列の認識手法に関する研究と、画像分類の精度を向上させるための手法の 1 つである画像データセットの拡張に関する研究について述べる。

* 連絡先：関西大学総合情報学部
〒569-1095 大阪府高槻市霊仙寺町 2-1-1
E-mail: mat@res.kutc.kansai-u.ac.jp

2.1 看板の文字認識に関する研究

情景画像中の看板に記述されている文字列を認識する研究は数多く行われている。Panhwar らは、機械学習を用いて自然景観画像の中にある看板画像に記述された文字情報に対し、認識が難しいウルドゥー語と英語を区別して認識する手法を提案した [6]。CNN を導入した MSER を用いて自然景観画像から看板画像を検出した後に、記述されている文字情報を同様に検出及び抽出し、HOG 特徴量及び SVM 分類器を用いて文字認識を行った。Kavati らは、OCR を用いて携帯端末のカメラに写っている看板の文字を認識し、英語からテグル語に翻訳してユーザに提示する旅行者向けの Web アプリケーションを開発した [4]。OCR は、手書き文字や印刷された文字などが写っている画像を解析し、その中に含まれている文字に相当するパターンを検出し、書かれている内容をテキストデータとして取り出す技術である。

2.2 画像データセットの拡張に関する研究

河野らは、GAN を用いた画像データセットの拡張手法を提案した [9]。GAN は生成モデルの 1 つであり、2 つのニューラルネットワークを交互に競合させて学習を進める事で、実在しないデータの生成や、存在するデータの特徴に沿った変換ができる手法である。これを用いた学習画像の生成によるデータセットの拡張を行った結果、分類モデルの正解率が上昇した。fujita らは、進化的画像処理を用いて指定された特徴を持つ画像を生成する手法を提案し、画像分類においてデータ拡張に適用することの有効性を検証した [2]。進化的画像処理は、最適化手法の一種である進化計算法を用いて、与えられた入力画像とその理想的な処理結果の画像である目標画像の組からなる学習画像セットによって、指定された画像変換を自動構築するものである。提案手法で水増画像を生成し、SVM 分類器を用いた上で、ランダムデータを用いた 2 種の分類問題によって精度を検討した結果、多くのケースの分類問題で正解率を向上させることができた。

2.3 本研究の位置づけ

2.1 節で述べた看板の文字認識を用いて、看板の文字をそのまま翻訳したとしても、ユーザはその店舗がどのような店舗であるかを理解できるとは限らない。そのため、本研究の目的を達成する場合において、ただ文字を翻訳してユーザの母国語に直すだけでは不十分である。

さらに、看板の文字が草書体のような認識が難しい書体である場合、看板から文字を正しく認識することが難しい。そのため、本研究では看板の文字認識を行わずに、ユーザへ店舗の情報を提示する。

2.2 節で述べた先行研究では、学習用画像の生成によるデータセットの拡張によって分類精度を向上させている。しかし、GAN を適用した水増し処理を行うには多くの学習用画像が必要となり、そのために全国数多にある各店舗の看板画像を用意することは容易ではない。進化的画像処理を用いた水増し処理においても、各店舗の看板画像ごとに特徴空間中の生成したい目標特徴点を指定する画像処理を行う必要があるため、多クラス分類において数多くある各店舗の看板画像ごとに行うには多大なコストがかかる。そのため本稿では、データセットの拡張を行わずに分類精度を向上させる手法を提案する。

3 看板画像を用いた店舗情報検索システム

本章では、本研究において基礎となる、看板画像から店舗の情報を取得するシステムについて述べる。北村らは、携帯端末のカメラを用いて看板から店舗の情報を取得するシステムを開発している。ユーザが目の前にある店舗の情報を取得する際、まず始めに、携帯端末のカメラを通して店舗の看板を写す。次に、カメラに写っている画像を、機械学習を用いた看板領域の検出及び看板画像の分類モデルによって識別を行う API サーバに送信する。看板領域の検出には Tensorflow1.0[1] で実装した YOLOv2[7]、看板画像の分類には、VGG16[5] を用いて構築した分類モデルを使用しており、これらを用いてユーザから送られてきた画像に写っている看板がどの店舗のものであるかを識別する。その後、OpenStreetMap [3] に登録されている識別した店舗に紐付いた店舗情報を、AR を用いてユーザの携帯端末に重畳表示する。これらの工程を 500 ミリ秒ごとに行うことで、リアルタイムな店舗情報の提示を可能とする。看板認識の処理を行うマシンのスペックを表 1 に示す。API サーバの構築には、Python3.6.5 と WebAPI フレームワークである Falcon1.4.1*1 を用いている。このシステムでは店舗の看板を分類する際、文字列ではなく看板画像を用いるため、草書体のような認識が難しい書体で記述されている看板であっても分類が可能である [8]。

北村らのシステムでは、大阪府高槻市にある商店街の一部の店舗を対象としている。そのため、システムを適

*1 <https://falconframework.org/> (2022/2/1) 存在確認

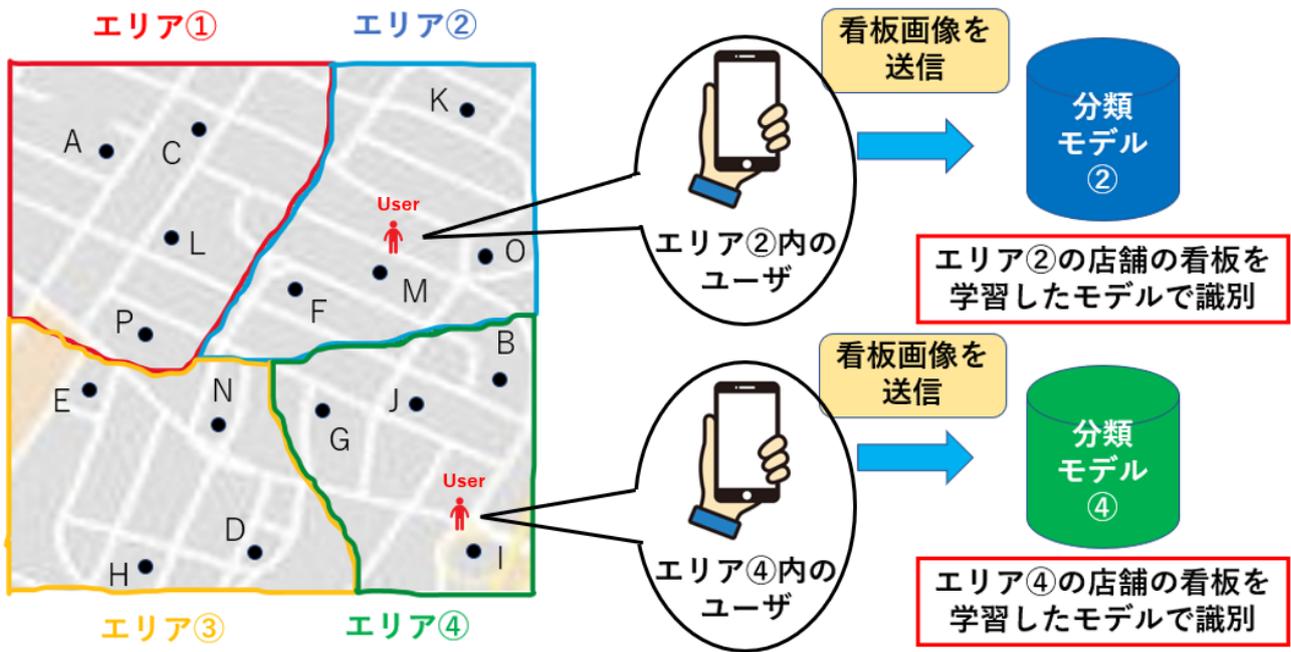


図1 ユーザの位置情報に基づく分類モデルの選定

表1 看板認識の処理を行うマシンのスペック

要素	スペック
CPU	Intel (R) Core (TM) i7-8700K @ 3.70 GHz
RAM	16 GB
GPU	NVIDIA GeForce GTX 1080 Ti
VRAM	12 GB
OS	Ubuntu 17.10

用できる店舗数が15店舗のみとなっており、特定の地域でしか使用できない。システムの実用性を上げるためには対象店舗数を増やす必要があるが、その場合は看板画像を分類する際に候補となる店舗数も同時に増加することになるため、それに伴って画像分類の精度が低下する懸念がある。そのため、システムの対象店舗を増加させることで適用範囲を拡大する場合、同時に画像分類の精度の低下を低減する必要がある。

4 提案手法

ユーザから送られてきた看板画像に対し、北村らのシステムでは対象店舗全ての看板画像を学習させた1つのモデルを使用して識別を行う。それに対して本手法では、予め特定の地域ごとに対象店舗の看板画像を学習さ

せた複数のモデルを構築し、その中からユーザが居る位置に応じて使用するモデルを選定して識別を行う。提案手法の概要を図1に示す。まず始めに、システムの対象地域を店舗間の地理的的近接性に基づいて分割し、それぞれのエリア内に位置する店舗の看板を学習させた画像分類モデルを、各エリアごとに構築する。ユーザから送られてきた看板画像をもとに店舗を識別する際、ユーザの携帯端末の位置情報に基づいてどのエリアにユーザが居るかを判定し、ユーザが居るエリア内に位置する店舗の看板を学習させた画像分類モデルを用いて識別を行う。これにより、各モデルにおいて店舗を識別する際に候補となる店舗数は、全ての店舗の看板画像を学習させたモデルと比べて大きく削減できるため、精度の向上が期待できる。加えて、システムの対象地域が拡大することで識別の対象となる店舗が増えた場合でも、システムが識別すべき店舗の数はエリアに属する店舗数に限定されるため、分類精度の低下を防ぐことが可能となる。

5 実装

提案手法の有用性を確認するため、以下の手順をもとに実装を行う。

1. 画像分類モデル構築のための看板画像データセットの作成
2. 地域の分割に用いる対象店舗の位置情報データセットの作成

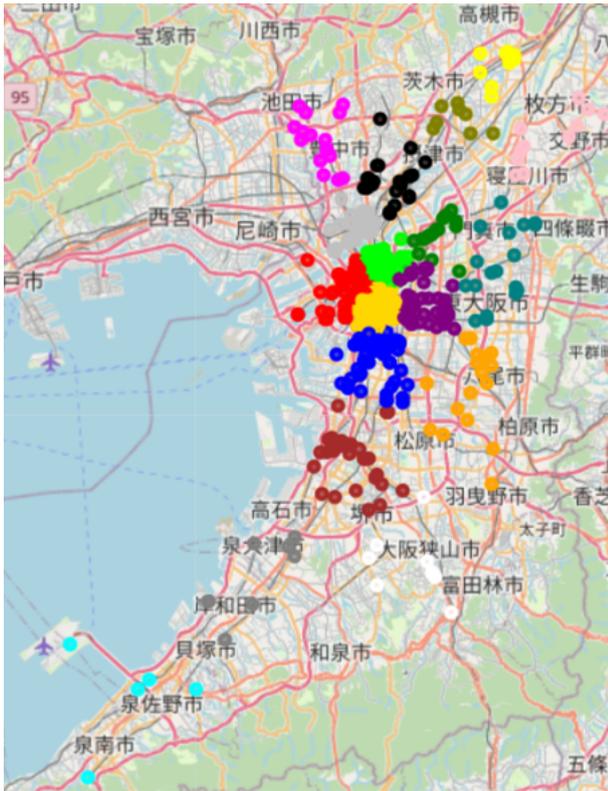


図2 対象地域の分割

3. 位置情報データセットをもとにした対象地域の分割
4. 分割した各エリアごとの画像分類モデルの構築

5.1 看板画像データセットの作成

北村らは看板画像データセットを作成する際、対象店舗の看板画像を様々な角度から撮影することで看板画像を収集している。しかし、店舗の看板を撮影して手動でデータセットを作成する場合、システムの対象店舗を増やす際に多大なコストがかかる。そのため、BeautifulSoup4によるスクレイピングを用いて、グルメサイトである食べログ*2から大阪府にある対象店舗の看板画像を自動で収集した。食べログでは地域を指定して店舗を検索することができ、大阪府は計44区域に分けられている。また、各店舗のページには、ユーザーによって投稿された店舗の外観の写真が多数掲載されている。本稿では対象地域から幅広く看板画像を収集するため、各区域ごとに人気ランキングページの上位から指定店舗数分の外観画像を全て収集した。

スクレイピングによって収集した外観画像に対し、

*2 <https://tabelog.com/> (2022/1/31 存在確認)

表2 各エリアの点の色と店舗数

エリア名	点の色	店舗数	エリア名	点の色	店舗数
エリア1	red	95	エリア10	orange	17
エリア2	brown	35	エリア11	white	9
エリア3	green	18	エリア12	pink	11
エリア4	blue	101	エリア13	silver	59
エリア5	yellow	20	エリア14	lime	195
エリア6	gray	7	エリア15	teal	18
エリア7	black	24	エリア16	aqua	5
エリア8	purple	95	エリア17	fuchsia	16
エリア9	gold	197	エリア18	olive	8

YOLOv2を用いて看板の物体検出を行った。北村らの物体検出モデルは、店舗の分類は行わず、看板の検出のみに特化しているため、様々な店舗の看板を検出できる汎用性を有する。このモデルを利用して信頼度が80%以上である看板領域を切り出し、切り出した看板画像が20枚以上存在している計930店舗を本システムの対象店舗として収集した。収集した各店舗の看板画像の60%をトレーニングデータ、20%をバリデーションデータ、20%をテストデータとしてデータセットを作成した。

5.2 対象店舗の位置情報データセットの作成

本手法では、前節でシステムの対象とした店舗同士の地理的近接性に基づいて対象地域の分割を行うため、前節で述べたスクレイピングを用いて各店舗のページに掲載されている住所を収集した。収集した住所に対し、対象店舗の居所の可視化及び電子的な処理を行うため、ジオコーディングを適用して各店舗の住所を緯度及び経度に変更した。ジオコーディングとは、住所、地名、郵便番号などが示す場所に対して地理座標を付与する処理を指す。本稿では、WebサービスであるGeocoding API*3を用いてジオコーディングを行った。このサービスでは、クエリとして住所やランドマーク名を入力して検索を行うことで、対応するGPS座標を表示できる。この変換作業をシステムが対象とする930店舗分全て手動で行う場合、多大な時間と労力を要するため、Selenium*4を用いて検索クエリの入力と検索の実行を自動で行った。その後、BeautifulSoup4を用いたスクレイピングによって検索結果から自動で対象店舗の緯度及び経度を収集し、対象店舗の位置情報データセットを作成した。

*3 <https://www.geocoding.jp/api/> (2022/1/31 存在確認)

*4 <https://www.selenium.dev/> (2022/1/31 存在確認)

表3 モデルの評価

モデル	店舗数	正解率	適合率	再現率	F 値
全店舗	930	0.799	0.714	0.625	0.667
エリア1	95	0.858	0.820	0.771	0.794
エリア2	35	0.839	0.854	0.819	0.836
エリア3	18	0.871	0.887	0.847	0.867
エリア4	101	0.821	0.826	0.772	0.798
エリア5	20	0.903	0.900	0.865	0.883
エリア6	7	0.911	0.923	0.907	0.915
エリア7	24	0.926	0.931	0.909	0.920
エリア8	95	0.825	0.826	0.785	0.805
エリア9	197	0.836	0.835	0.788	0.811
エリア10	17	0.858	0.897	0.846	0.871
エリア11	9	0.886	0.939	0.872	0.904
エリア12	11	0.94	0.959	0.919	0.939
エリア13	59	0.876	0.855	0.822	0.838
エリア14	195	0.824	0.805	0.765	0.785
エリア15	18	0.848	0.802	0.809	0.805
エリア16	5	1.0	1.0	1.0	1.0
エリア17	16	0.943	0.943	0.930	0.936
エリア18	8	0.857	0.896	0.846	0.870

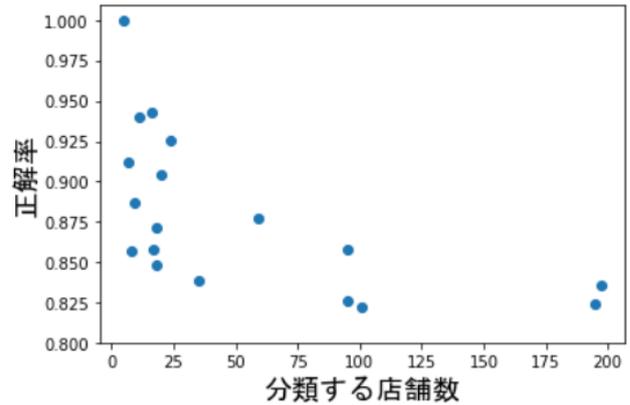


図3 分類する店舗数と正解率の関係

5.3 対象地域の分割

対象地域内に位置する対象店舗は、規則性に基づいて立地しているわけではないため、対象店舗間の地理的接近性をもとに対象地域を分割し、各エリアごとに大きさを変える必要がある。そのため、前節で作成した位置情報データセットをもとに、k平均法によるクラスタリングを行った。クラス数は各エリア内の店舗数と分類精度のトレードオフを考慮して18クラスに設定している。その後、図2のように地理的接近性に基づいて対象地域を分割した。各点は店舗の所在地を示しており、各点の色は店舗が分割したエリアのどこに位置しているかを示す。各エリアに対応する点の色と、各エリアに属する店舗の数を表2に示す。

各エリアに属する店舗数は様々であり、最も店舗数が多いエリア9は197店舗であったのに対し、最も店舗数が少ないエリア16は5店舗であった。都市部に近い店舗ほど、システムの対象とするために必要な量の看板画像を集めることができた店舗が多かったため、都市部に近いエリアであるほど多くの対象店舗が属していた。

5.4 看板の画像分類

5.1節で作成した看板画像データセットをもとに、CNNモデルであるVGG16を用いて画像分類を行う。VGG16とは、1000クラスに分類された130万枚の画像を学習させた、13層の畳み込み層と3層の全結合層の計16層からなるモデルである。本稿では北村の手法と同様に、全結合層を取り除き、新たに全結合層を追加する転移学習を行っており、学習に用いる画像が少ない店舗であって短時間で高精度なモデルを構築できる。

6 看板分類モデルの評価と考察

6.1 評価

提案手法の有用性を確認するため、対象地域にある全ての930店舗の看板画像を学習させたモデルと、対象地域を18区域に分割し、エリアごとに看板画像を学習させて構築したモデルの精度を比較した。各モデルの画像サイズを224×224、バッチサイズを16、学習率を 10^{-4} 、慣性を0.9とした。エポック数は、930店舗を分類するモデルは2000とし、エリアごとに看板画像を学習させて店舗を分類するモデルは、各モデルのクラス数に合わせて50-400とした。損失関数には、各モデルのクラス数の交差エントロピーを用いた。

評価指標には正解率及び適合率、再現率、F値を使用し、テストデータを用いて算出した。各モデルの分類する店舗数及び計算結果を表3に示す。全ての対象店舗の看板を学習させたモデルの正解率は0.799であったのに対し、対象地域を分割してエリアごとに構築した18区域のモデルの正解率の平均は0.879であった。また、エリアごとに構築した全てのモデルにおいて、全ての対象

店舗の看板を学習させたモデルと比較して正解率が高かった。以上の結果から、提案手法を用いることによる分類精度の向上が確認できた。

6.2 考察

対象地域を分割してエリアごとに構築した各モデルの、分類する店舗数と正解率の関係を図3に示す。分類する店舗数が多いモデルほど正解率が低い傾向を確認でき、最も正解率が高いモデルはエリア16の1.0であるのに対し、最も正解率が低いモデルはエリア4の0.821であり、エリアによってモデルの精度が大きく異なっていた。分割されたエリアの1つに店舗が集中した場合、その地域の店舗の看板画像を学習させたモデルの精度が大きく低下する懸念がある。そのため、各エリアごとの店舗数を可能な限り均一にすることで、各モデルの分類精度の差を低減するように地域を分割する必要がある。

エリアを細分化して1つのエリア辺りの店舗数を削減することで、各モデルの更なる精度向上が期待できる。しかしそれに伴って構築するモデル数も同時に増えることになり、構築するモデルが増えるほど多大なコストがかかる。そのため、1つのエリアあたりの店舗数と構築するモデル数のトレードオフを考慮して地域を分割する手法を検討する必要がある。

7 課題と展望

データセットとなる看板画像を食べログから収集した際、20枚以上収集できなかった店舗は、画像分類の精度を十分に確保できないためにシステムの対象外としている。しかし、大阪府には食べログに載っている店舗だけでも6万店舗以上あり、システムが適用可能な店舗はごく一部に留まっている。今後システムの適用範囲を広げる場合、看板画像を収集する手法や看板画像を学習させる手法を検討する必要がある。

看板画像データセットの中には、建物の壁面の一部や張り紙のような、学習に不要な画像が多く混ざっていた。これらの画像の混入がシステムの精度に影響を与えていると考えられる。そのため、更なる分類精度の向上のためにはデータセットに含まれる不要なデータを取り除くか、看板の物体検出の精度を向上させる必要がある。

8 終わりに

本稿では、看板画像を用いた店舗情報検索システムにおいて、画像分類モデルの精度を向上させる手法を提案し、その有用性を確認した。今後は、6.2節で述べた考察

をもとに、地域の分割手法の改良を行う。加えて、ユーザの位置情報に基づいて自動で識別する際に使用するモデルを選定できるようにシステムの改良を行い、実世界で看板画像の識別精度が向上するかを検証する。

参考文献

- [1] Abadi, M. et al.: TensorFlow: A System for Machine Learning, *12th USENIX Symposium on Operating Systems Design and Implementation*, pp. 265–283 (2016).
- [2] Fujita, K. et al.: Data Augmentation using Evolutionary Image Processing, *2018 Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–6 (2018).
- [3] Haklay, M. and Weber, P.: OpenStreetMap: User-Generated Street Maps, *IEEE Pervasive Computing*, pp. 12–18 (2008).
- [4] Kavati, I. et al.: Signboard text translator: A guide to tourist, *International Journal of Electrical and Computer Engineering*, Vol. 7, No. 5, pp. 2496–2501 (2017).
- [5] Liu, S. and Deng, W.: Very deep convolutional neural network based image classification using small training sample size, *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 730–734 (2015).
- [6] Panhwar, M. A. et al.: Signboard Detection and Text Recognition Using Artificial Neural Networks, *IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, pp. 16–19 (2019).
- [7] Redmon, J. and Farhadi, A.: YOLO9000: Better, Faster, Stronger, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517–6525 (2017).
- [8] 北村茂生, 松下光範: オンサイト検索: 携帯端末を用いた看板画像からの店舗アクセス手法, *電子情報通信学会論文誌*, Vol. J103-D, No. 5, pp. 364–372 (2020).
- [9] 河野曜平, 川本一彦: GANを用いたデータ拡張, *情報処理学会研究報告*, Vol. 2017-CVIM-207, No. 14, 情報処理学会, pp. 1–5 (2017).
- [10] 野村松信ほか: 背景色情報を活用した看板内の文字列領域抽出法, *電気学会論文誌C*, Vol. 134, No. 1, pp. 121–130 (2014).

主題関連度を用いた課題達成のための対話の自動評価

Automatic Evaluation of Dialog for Task Achievement using Subject Relevance

萬徳 昭太 砂山 渡*
Shota Mantoku Wataru Sunayama

滋賀県立大学 工学部
School of Engineering, The University of Shiga Prefecture

Abstract:

対話の場面においては、時間の有益かつ効率的な利用が求められるが、対話を人手で評価することは容易ではなく、評価者の負担も大きい。そこで本研究では、課題達成のための対話の書き起こしデータについて、対話の中で用いられる主題に関する単語との関連度を用いて、各発話者の主題に関する発話量をもとに、対話の自動評価を行うシステムを提案する。

1 はじめに

対話の場面においては、時間の有益かつ効率的な利用が求められる。特に何らかの結論を出す必要がある対話においては、結論に結びつくアイデアを並べる発散フェーズと、挙げられたアイデアを整理していく収束フェーズとがあり、それぞれにおいて、結論に結びつく効果的な対話が行われることが望まれる。このような対話の場面においては、目的に沿った対話が行われているか、また対話に参加している全員が効果的な発言をしているかなどを、対話のリーダーシップを取る人が調整する必要がある。しかし、対話を人が客観的に評価することは容易ではなく、また評価者の負担も大きい。

そこで本研究では、課題達成のための対話の書き起こしデータについて、対話の中で用いられる主題に関する単語との関連度を用いて、各発話者の主題に関する発話量をもとに、対話の自動評価を行うシステムを提案する。これにより、対話の評価の自動化、ならびに対話の内容や方法の改善について、事後評価やリアルタイム評価などへの応用が期待できる。

2 関連研究

対話の場において、参加者の発言の自動評価を行い、音響的特徴と言語的特徴に基づく評価の分析結果をユーザにフィードバックをする研究 [1] や、前後の発言と兼ね合いを考慮しながら、時系列データとして発言内容

の説得力を評価する研究 [2] がある。本研究においては、対話の主題となる単語を設定した上で、発話に相当する対話中の各文の、主題との関連度をもとに対話全体を評価する。

議論能力に基づく Wikipedia における編集者間議論ページの分析を行う研究 [3] もあり、Wikipedia で行われている記事の修正や編集方針について、建設的な議論が行われることもあるが、編集者間の相互誹謗により無意味な論争が長く続く非効率な場合も見られる。そのため、建設的な議論の総量として、議論の主題と関わりをもとに対話の評価することが望まれる。

議論マップと呼ばれる議論を可視化するツールを作成する研究 [4] や、多人数が受講する講義における少人数で行う議論を対象として、他のグループの議論内容から活発さを反映した共起ネットワークを作成して提示する研究 [5] がある。本研究においては、対話の具体的な様子の理解ではなく、対話全体の総合的な評価を与える点数づけを行う。

3 主題関連度を用いた課題達成のための対話の自動評価システム

本章では、主題関連度を用いた課題達成のための対話の自動評価システムについて述べる。

3.1 本研究で用いる対話データ

本研究で用いる対話データは、指定の課題に対して新しい事業案を作成するために行われたもので、主に

*連絡先： 滋賀県立大学工学部 砂山 渡
〒 522-8533 滋賀県彦根市八坂町 2500
E-mail: sunayama.w@e.usp.ac.jp

表 1: 分析に用いた対話データと人手による評価点

グループ	人数	E1	E2	E3	評価点
A	4	○	○	-	8
B	4	◎	○	○	10
C	4	△	-	-	5
D	4	-	-	-	6
E	4	◎	-	-	8
F	4	△	-	△	4
G	4	◎	-	-	8
H	4	◎	○	-	9
I	5	△	△	△	3
J	4	◎	○	○	10
K	4	△	△	-	4
L	4	◎	○	○	10
M	3	○	○	○	9
N	4	○	○	○	9
O	4	-	-	△	5
P	4	-	-	-	6
Q	3	○	○	○	9
R	4	-	-	-	6
S	4	◎	-	○	9
T	4	○	○	-	8

4名によって20分程度のオンライン対話が行われた音声データをテキストに書き起こしたものとなっている(表1)。

本データは、ある人事関連サービスを行う会社から研究用にご提供いただいたもので、各対話データには、同社内で人手による、以下の評価が行われている。本研究では、その評価に基づいて、◎を4点、○を3点、-を2点、△を1点として、点数化した評価点を用いる。

- E1: 総合評価
 - ◎大変優れている
 - ○優れている
 - -ふつう
 - △不足している
- E2: 活動量
 - ○目立って多かった
 - -ふつう
 - △目立って少なかった
- E3: 建設的議論
 - ○目立って見られた
 - -ふつう
 - △目立って少なかった

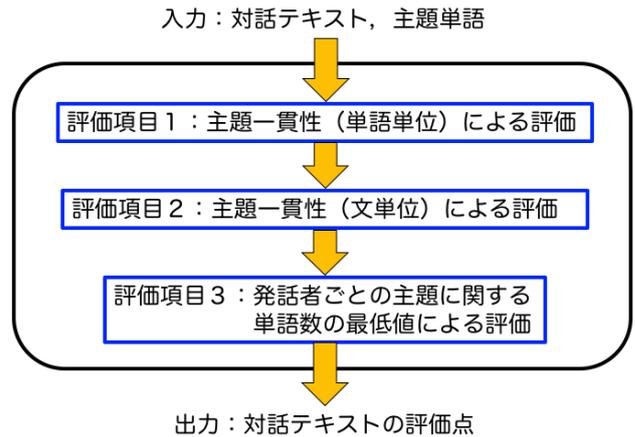


図 1: 主題関連度を用いた課題達成のための対話の自動評価システムの構成

3.2 主題関連度を用いた課題達成のための対話の自動評価システムの構成

主題関連度を用いた課題達成のための対話の自動評価システムの構成図を図1に示す。本システムには、評価対象となる対話テキストと、同テキストの主題を表す単語を入力する。システムは、入力された対話テキストに対して、入力された主題単語に関する3つの項目についてテキストの評価を行い、その評価結果を点数として出力する。

3.3 入力: 主題単語の設定

本研究で用いる対話テキストに対して、設定した主題単語を表2に示す。これらの主題単語は、各対話テキストにおける頻度が高い名詞の中から、「対話のテーマ名に含まれる単語」または、対話の内容と関連のある単語として設定した。

3.4 評価項目1: 主題一貫性(単語単位)による評価

主題一貫性評価(単語単位)による評価値を、式(1)で与える。ただし、式中の主題に関係する単語は、文献[6]による主題関連語としており、この値はフリーソフトのテキスト分析ツールTETDM[7]によって取得する。

$$\text{主題一貫性評価(単語)} = \frac{\text{主題単語または主題関連語の数}}{\text{単語の総数}} \quad (1)$$

表 2: 対話テキストに設定した主題単語

グループ	主題単語
A	盆栽 他2語
B	工芸品 他2語
C	子ども 他1語
D	生徒 他1語
E	オンライン 他3語
F	環境負荷 他1語
G	廃棄 他5語
H	インドネシア 他4語
I	インターシップ 他1語
J	教育 他1語
K	健康 他1語
L	中国 他5語
M	観光 他3語
N	人材 他4語
O	ボランティア 他3語
P	産業 他3語
Q	オンライン 他4語
R	能力 他4語
S	会社 他3語
T	仕事 他3語

表 3: 主題一貫性評価 (単語単位) の評価点

評価点	主題一貫性 (単語) [%]
4	40 % 以上
3	30 % 以上
2	20 % 以上
1	20 % 未満

主題関連語は、対話データの先頭の文から順に、主題単語または主題関連語が出現する文に出現する単語として逐次的に与えられる。また、主題一貫性評価 (単語単位) の評価点を、表 3 のように与える。

本評価点を設定した理由は、課題達成に向けては、主題に関する単語を利用することが望まれる、と考えたことによる。

3.5 評価項目 2 : 主題一貫性 (文単位) による評価

主題一貫性評価 (文単位) による評価値を、式 (2) で与える。ただし、式中の主題に関係する文は、文献 [8] による主題関連文としており、この値はフリーソフトのテキスト分析ツール TETDM によって取得する。

$$\text{主題一貫性評価 (文)} = \frac{\text{主題に関係する文の数}}{\text{文の総数}} \quad (2)$$

表 4: 主題一貫性評価 (文単位) の評価点

評価点	主題一貫性 (文) [%]
4	20 % 以上
3	15 % 以上
2	10 % 以上
1	10 % 未満

表 5: 発話者ごとの主題に関する単語数の最低値の評価点

評価点	主題に関する単語数の最低値
4	15 個以上
3	12 個以上
2	10 個以上
1	10 個未満

主題関連文は、いずれかの主題単語または主題関連語が出現する文として与えられる。また主題関連語は、対話データの先頭の文から順に、主題単語または主題関連語が出現する文に出現する単語として逐次的に与えられる。また、主題一貫性評価 (文単位) の評価点を、表 4 のように与える。

本評価点を設定した理由は、課題達成に向けては、基本的に主題に関する話をする事が望まれる、と考えたことによる。

3.6 評価項目 3 : 発話者ごとの主題に関する単語数の最低値による評価

本評価項目の値は、各参加者の主題単語または主題関連語を使用した回数を調べた上で、その参加者内での最低値として定義する。

主題関連語は、他の評価項目で定義した単語と同一のものをを用いる。また、主題に関する単語数の最低値による評価点を、表 5 のように与える。

本評価点を設定した理由は、課題達成に向けては、全員が主題に関する意見を述べる必要がある、と考えたことによる。

3.7 出力 : 対話テキストの評価点

前節までで述べた 3 つの評価項目による評価点を組み合わせる値を、対話テキストの評価点として出力する。表 6 に、本研究で用いた対話テキストの評価点を示す。

表 6: 各評価項目による対話テキストの評価点

グループ	評価項目 1	評価項目 2	評価項目 3
A	3	4	2
B	4	4	4
C	2	4	2
D	2	4	1
E	3	4	2
F	2	1	1
G	3	4	4
H	3	3	1
I	1	2	1
J	4	4	3
K	2	2	1
L	4	4	4
M	3	4	4
N	4	3	3
O	1	3	3
P	2	3	1
Q	3	4	4
R	2	1	1
S	3	4	3
T	4	4	3

4 対話の自動評価システムの評価

本章では、3章の「主題一貫性（単語単位）」「主題一貫性（文単位）」「発話者ごとの主題に関する単語の数の最低値」の3つの評価項目の評価の妥当性とその組み合わせ方について検証した結果について述べる。

4.1 3つの項目の組み合わせによる点数化の方法の妥当性の検討

表7に、人手による3つの評価項目による評価点数と、システムによる3つの評価項目による評価点数との相関係数を示す。

E1からE3の人による各評価項目について、システムの評価項目1の「主題一貫性（単語単位）」、または、評価項目1×評価項目2+評価項目3が、およそ0.8以上の高い相関係数の値を示した。特に、人手によるE1+E2+E3の評価項目と、システムによる評価項目1×評価項目2+評価項目3との相関が0.90と最も高い値となった。このことから、システムの評価によって、人手に近い評価が得られたことがわかる。

人手による評価項目ごとに見ると、全体的に評価項目1との相関が高い傾向にあった。また、E1においては、評価項目1×評価項目2+評価項目3との相関が最も高くなった。このことから、人手による評価においては、具体的な対話の内容に関わる評価では、単語単位での主題一貫性が重要視されていること、また対話

の結論に対しての総合評価としては、対話参加者全員の関与が必要な要素になっていると考えられる。

E1+E2+E3の人手の評価の合計に対しては、評価項目1との相関と、評価項目1×評価項目2+評価項目3との相関がほぼ同じ値となったが、先の考察の結果を踏まえると、3つの評価項目を組み合わせた、評価項目1×評価項目2+評価項目3によって評価していくことが望ましいと考えている。

5 おわりに

本研究では、課題達成のための対話の書き起こしデータについて、対話の中で用いられる主題に関係する単語との関連度を用いた3つの評価項目を提案し、対話の自動評価を行うシステムを提案した。提案する評価項目とその組み合わせにより、対話の自動評価を行える可能性を検証した。

今後は、対話の事後評価をもとにした、対話内容の改善を図るためのフィードバックを行うことや、リアルタイムの対話の中で、自動的に主題単語を設定した上で、リアルタイムに対話の評価を行うシステムなどへの応用を検討している。

謝辞

本研究の対話データは、(株)リクルートマネジメントソリューションズ組織行動研究所の今城研究員より、研究用データとしてご提供いただいたものです。ここに記して感謝申し上げます。

参考文献

- [1] 岡本康佑, 松原茂樹, 長尾確: 会議における発言の音響的特徴と言語的特徴に基づく自動評価, 情報処理学会第78回全国大会, pp.521 - 522 (2016)
- [2] 澤田慎太郎, 中川智皓, 新谷篤彦, 井之上直也: 対話的議論の自動評価に向けたディベートデータセットの構築, 言語処理学会第26回年次大会発表論文集, pp.708 - 711 (2020)
- [3] 朱成敏, 武田英明: 議論能力に基づく Wikipedia における編集者間議論ページの分析, 人工知能学会第26回全国大会, 1C2-R-5-1 (2012)
- [4] 趙雲超, 松村真宏, 谷内田正彦: 音声認識された議事録からの議論マップ自動生成, 人工知能学会第20回全国大会, 3C3-5 (2006)

表 7: 人とシステムの評価点の相関係数

評価 (人) / 評価 (システム)	1 (一貫性・単語)	2 (一貫性・文)	3 (主題最低発話)	1+2+3	1×2+3	1×3+2	2×3+1	1×2×3
E1 (総合評価)	0.78	0.64	0.62	0.78	0.82	0.75	0.72	0.74
E2 (活動量)	0.79	0.52	0.59	0.73	0.75	0.71	0.66	0.69
E3 (建設的議論)	0.78	0.57	0.64	0.77	0.78	0.79	0.73	0.77
E1+E2+E3	0.89	0.67	0.70	0.87	0.90	0.85	0.80	0.83

- [5] 村岡泰成, 石川誠彬, 尾澤重知, 江木啓訓: 議論の活発さを反映した共起ネットワーク図の提示による効果の検討, 情報処理学会第 110 回グループウェアとネットワークサービス研究会, No.8 (2020)
- [6] 砂山渡, 谷川信弘: テキストの主題に基づく一貫性評価と結論抽出への応用, 日本知能情報ファジィ学会誌, Vol.23, No.5, pp.727 – 738 (2011)
- [7] 砂山渡, 高間康史, 徳永秀和, 串間宗夫, 西村和則, 松下光範, 北村侑也: 統合環境 TETDM を用いた社会実践, 人工知能学会論文誌, Vol.32, No.1, NFC-A, pp.1 – 12 (2017)
- [8] 西原陽子, 佐藤圭太, 砂山渡: 光と影を用いたテキストのテーマ関連度の可視化, 人工知能学会論文誌, Vol.24, No.6, pp.480 – 488 (2009)