

音声対話およびTwitterにおけるユーザのパーソナリティ自動推定

藤倉将平 小川義人 菊池英明
早稲田大学人間科学部 早稲田大学人間科学学術院 早稲田大学人間科学学術院
Ospirall@asagi.waseda.jp stream@toki.waseda.jp kikuchi@waseda.jp

1 はじめに

1.1 本研究の目的

音声対話を通じて人間の特定のタスクを請け負う新しいサービスとしてAppleのSiriやNTTdocomoのコンシェルジュサービスのようなアシスタントサービスが台頭してきている。いずれも特徴的であるのは、人を模したエージェントを用意することでユーザが難しく面倒な操作を行う必要性を減らしかつ、ユーザのインタラクション欲求を高めていることである。そしてまたこれらのエージェントを代表する、人間とインタラクションをはかるシステムのニーズは今後増えていくと共に、エージェントと人間とのインタラクションをより良いものにしていくことが期待されるだろう。筆者は対話システムがユーザのパーソナリティを自動的に推定し、その情報を活用した応答を返すことによってユーザが対話を継続したいと思えるようなシステムを提案する研究を行っている。本稿では提案システム実現の為の手法として、ユーザのパーソナリティを自動的に推定するシステムを構築する。本研究においてはシステムとユーザの対話を行うプラットフォームとして、音声対話とTwitterの2つを用いた。Twitterを採用した理由は、音声対話のみだけでは得る事が難しい情報が収集できる可能性があると考えられるためである。またユーザは音声対話におけるパーソナリティとTwitter上でのパーソナリティの2つを持つと仮定する。

1.2 パーソナリティの定義

本研究ではユーザのパーソナリティをBigFive尺度[1]によって表現する。ユーザは外向性、神経症傾向、開放性、調和性、誠実性をそれぞれ0~50の点数で表される。またBigFive尺度のそれぞれに対して正の因子負荷量が高いものをポジティブ、負の因子負荷量が高いものはネガティブと呼ぶ。例えば“陽気な”、“話し好き”などは外向性においてポジティブであり“人嫌い”、“地味な”などは外向性においてネガティブである。

2 ユーザのパーソナリティ自動推定

2.1 推定方法

パーソナリティ推定システムはユーザが過去に発言したテキストデータから、その人がBigFive尺度のそれぞれの尺度をどの程度持つのか数値化することにより、ユーザのパーソナリティを自動的に推定する。ユーザの発言に対してBigFive尺度のそれぞれにおいてポジティブとネガティブとで二値分類を行う。分類の結果はポジティブなものと判断された場合に1、ネガティブなものと判断された場合は0を返すものとする。それぞれの尺度において分類された直近の50個までのデータを足し合わせた0~50の得点をそのユーザの持つパーソナリティとする。

2.2 分類器の構築

分類器は半教師多項式ナイーブベイズ分類器[2]を使う。ラベル有り学習データセットは後述する調査を行う事によって収集した。ラベルなしデータセットはTwitterにおけるツイートをを用いる。素性として形態素を扱うため文を形態素解析器であるMeCab[3]に渡すことで形態素に変換する。

2.3 ラベル付きデータセット収集のためのアンケート調査 1

ラベル付きデータセットを収集するため,CGIを通して自由回答のアンケート調査を行った.被験者はあるパーソナリティを保持している人がどのような発言を行うか意識して文を作成した.被験者は10名となった.

表 1: 交差検定の結果

BigFive尺度	ラベル	分類失敗文数	総文数	分類成功率
外向性	ポジティブ	26	115	77.39%
	ネガティブ	29	114	74.56%
神経症傾向	ポジティブ	32	114	71.93%
	ネガティブ	23	117	80.34%
開放性	ポジティブ	46	117	60.68%
	ネガティブ	29	116	75.0%
調和性	ポジティブ	26	113	76.99%
	ネガティブ	32	113	71.68%
誠実性	ポジティブ	38	115	66.96%
	ネガティブ	32	115	72.17%

表1はアンケート調査によって集められたデータセットの交差検定の結果を表している.フォールド数は5とした.分類成功率は10項目を平均して約72.77%となった.

2.4 ラベル付きデータセット収集のためのラベル付け調査

この実験では2.3節のアンケート回答者の意識にのぼらないような発言表現を収集することを目的とした.CGIを通して被験者に対話コーパスの書起しとTwitterから集めた「ツイート」をランダムに提示し,被験者はその文から発言者がどのようなパーソナリティをもっているか判断し,ラベル付けを行う.被験者は10名となった.

表 2: 交差検定の結果

BigFive尺度	ラベル	分類失敗文数	総文数	分類成功率
外向性	ポジティブ	82	234	64.96%
	ネガティブ	45	96	53.13%
神経症傾向	ポジティブ	60	203	70.44%
	ネガティブ	51	145	64.83%
開放性	ポジティブ	66	131	49.62%
	ネガティブ	48	114	57.89%
調和性	ポジティブ	25	95	73.68%
	ネガティブ	34	86	60.47%
誠実性	ポジティブ	46	91	49.45%
	ネガティブ	77	251	69.32%

表2はラベル付け調査によって集められたデータセットの交差検定の結果を表している。フォールド数は5とした。分類成功率は10項目を平均して約61.38%となった。

2.5 構築された分類器の評価実験

表 3: 交差検定の結果

BigFive尺度	ラベル	分類失敗文数	総文数	分類成功率
外向性	ポジティブ	136	464	70.69%
	ネガティブ	81	210	61.43%
神経症傾向	ポジティブ	90	317	71.61%
	ネガティブ	82	262	68.7%
開放性	ポジティブ	98	248	60.48%
	ネガティブ	75	230	67.39%
調和性	ポジティブ	57	208	72.6%
	ネガティブ	74	199	62.81%
誠実性	ポジティブ	77	206	62.62%
	ネガティブ	117	366	68.03%

表3は2.3と2.4節で収集したデータを用いた分類器の交差検定の結果を表している。フォールド数は5とした。分類成功率は10項目を平均して約66.66%となった。

3 質問用紙法とパーソナリティ推定システムの比較

従来のBigFiveを測定するために使用されてきた質問用紙によるユーザのパーソナリティと、本研究で使用するパーソナリティ推定システムによって推定されたユーザのパーソナリティを比較する。質問用紙による質問内容は[4]を参考にした。被験者は大学生11名となった。

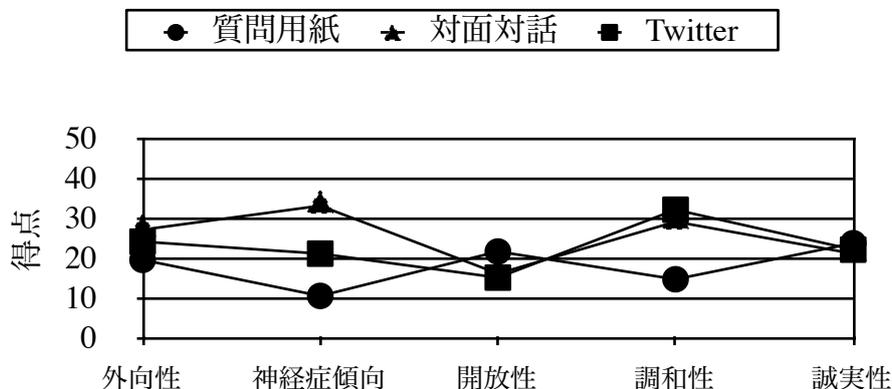


図 1: ある被験者のパーソナリティ推定結果

図1はある被験者のパーソナリティを質問用紙法とパーソナリティ推定システムで比較したものを表している。質問用紙での得点を正解データとしたとき、この被験者は外向性、開放性、誠実性についてはよく推定されている。

■ 質問用紙 - 音声対話 ■ 質問用紙 - Twitter ■ 音声対話 - Twitter

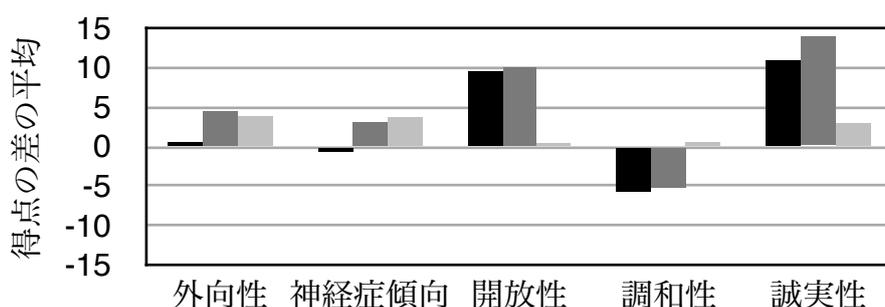


図2: それぞれの手法における全被験者のパーソナリティの得点の差の平均

図2はそれぞれの手法における全被験者のパーソナリティの得点の差の平均を表している。調和性は質問用紙では低く得点されているが、音声対話とTwitterにおけるパーソナリティ推定システムによると高い得点を持つと推定されている傾向が強かった。開放性と誠実性においては、推定された結果は質問用紙の得点より低く推定されている傾向が強かった。これは対話を行う際には、「無駄」、「めんどい」のような開放性、誠実性においてネガティブに捉えられる表現が多かったこと、また調和性において「~ください」のようなポジティブに捉えられる表現が多かったことが考えられる。

4 まとめ

本研究ではパーソナリティを自動的に推定するシステムを構築し、質問用紙の手法とのBigFive尺度における得点の違いを述べた。今後の課題として2.5節で示した分類器の性能が約66.66%に留まったが、2.4節で集めたラベル付きデータセットの中にはポジティブとネガティブのどちらもつかないような表現が含まれていた。これらの妥当性を検証することで分類の成功率を高める事が出来ると考えられる。また、どれだけの履歴を保持する事によってユーザのパーソナリティをより正確に推定できるようになるのか調査する必要があると考えられる。

今後の予定としてパーソナリティ推定システムを利用し、対話システムがユーザのパーソナリティ情報を用いた対話表現を行う事によって、ユーザの対話継続欲求を向上させる研究を進める。

参考文献

- [1] 齊藤崇子,中村知靖,遠藤利彦,横山まどか ”性格特性用語を用いたBigFive尺度の標準化”,九州大学心理学研究 11 2 11,p135-144 (2001).
- [2] Jiang Su, Jelber Sayyad Shirab, Stan Matwin, ”Large Scale Text Classification using Semi-supervised Multinomial Naive Bayes”, ICML, Washington, USA, 2011.
- [3] MeCab <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html> version 0.993.
- [4] 堀洋道, ”心理尺度集I”, 山本眞理子 (編), サイエンス社, 東京, 2001.