



複数種データを用いたAIによる感情分析

Emotion Analysis AI by Several Types of Data

県立福島高等学校
2年 二階堂 遼
2年 岡部 吹
2年 半澤 慶也

<Abstract>

According to the National Police of Japan homepage, the number of suicides in Japan remains high. In 2020, it increased for the first time in 10 years. It seems to be that one of the causes is the change in social activities because of the spread of COVID-19. We considered that emotion analysis AI might contribute to the decrease in suicides by supporting counselors. Most of the existing emotional analysis services use only one analysis. However, can using the data on facial expression, voice recognition, and natural language processing, along with Python3, upgrade the accuracy in its analysis? We show the results of emotional detection with a video of a speech by the former prime minister Abe. We compared the emotions read from the facial expressions with general opinions and the analytical results of the existing service. Those two comparison results show a low rate of concordance. It cannot be said that the accuracy was upgraded. From now on, we are going to examine the discriminant model carefully and think about how to correct it.

1. 感情分析AIとは

表面的な人間の感情、気持ちの状態や変化を読み取るAIのことを言う。
分析の対象は文章(自然言語処理)、顔の表情、声など様々である。
やがて、人間のあらゆる言動、瞳孔の動き、脈拍の変化、発汗の様子、脳波のパターンなどを分析できるようになると言われている。

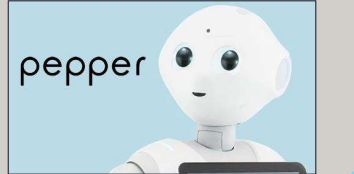


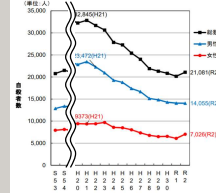
図1. ソフトバンクのPepper

2. 研究背景

表1-1. 警察庁HP「令和二年度中における自殺の状況 資料」

全国合計	青森	岩手	宮城	秋田	山形	福島	東京	...	神奈川	...	大阪	...	沖縄
20830	292	198	418	184	211	356	2293	...	1181	...	1340	...	238

表1-2. 自殺者数の年次推移



感情分析AIを用い、全国に多い自殺者数を減らせないと考えた。

具体的には、精神を病んだ方々が受診するカウンセラーの業務をサポートすることができるソフトウェアの開発を行う。

AIが受診者の精神状態を推測し、第三者的な分析結果を提供する。

3. 仮定 複数種のデータを一度に用いた感情分析サービスは見つからなかった。

表2. 現状の感情分析サービスと私たちの研究比較

例	私たち	企業
トレーニングデータ数	3	1
データの種類	表情・声・自然言語処理	(確認した限り)
合計データ数	× 12	× 12

企業よりも得られるトレーニング・テスト用データは少ないが、コミュニケーションにおいて声だけを聴くより顔を見ながら話した方が気持ち伝わりやすいように、表情や声の情報を同時に用いることで精度の向上が望めるのではないかと考えた。

4. 研究意義

精神状態の異変の自覚を促す複数手法を用いた感情分析AIの開発。

5. リサーチクエスト

複数種のデータを用いたトレーニング・分析によって精度は向上するか。実際、実用的なサービスと言えるのか。

7. 実験結果 使用動画: 安倍前総理大臣のスピーチ

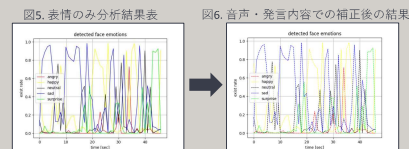
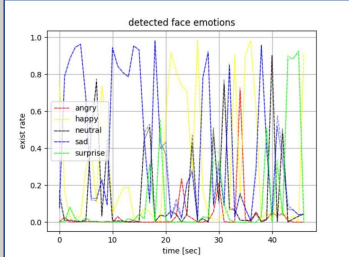


図5・図6の比較 (透過して重ね)



8. 考察

補正前後で比較的参数変動が大きかった29秒における結果を示す

	補正前 (割合)	補正後 (割合)
Angry	0.055435088	0.05705038
Happy	0.012243598	0.013027488
Neutral	0.51129967	0.459308554
Sad	0.09592838	0.087847257
Surprise	0.3250934	0.382766321

図7. 29秒の顔画像

生徒へのアンケート (一般意見の検証)

表4. 生徒へのアンケート結果

	1番存在する (%)	2番目に存在する (%)
Angry	25.5	38.3
Happy	2.1	4.2
Neutral	66	23.4
Sad	6.4	27.7
Surprise	0	6.4

一般人は互いの表情についてどの感情が優勢だと考えるのかを検証するため、生徒へのアンケートを行った。(回答者数47人)
1番存在すると思う感情はNeutral、2番目に存在すると思う感情はAngry、だという結果になった。

既存サービスでの検証

	結果 (割合)
Angry	0.153
Happy	0.05
Neutral	0.229
Sad	0.388
Surprise	0.181

既存サービスでの判定との差異を確かめるため、企業User localが提供する表情推定(感情認識)AIで判定を行なった。
存在確率1位はSad、2位はNeutralという結果になった。

- 補正前後で存在確率の高い感情に変化がない
- 上記2つの検証において、一致率が低い

精度未向上

6. 処理の流れ

FACE

- 表情判別モデルを作成
→畳み込みニューラルネットワーク (CNN) を使用
- 判別する人物の顔を定義
→目的人物以外の顔が映った時に誤認する可能性がある
- 各秒数の顔画像に対して感情分析
→1で作成したモデルで分析



図2. CNNの仕組み

VOICE

- 動画から音声を作成
- 4秒ごとの音声に切り分ける
→Web Empath APIでは5秒以上の音声を分析できない
- 各音声をAPIで感情分析



図3. Empathロゴ

SENTENCE

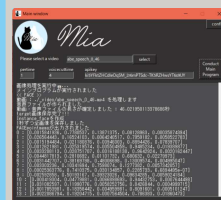
- 音声のノーマライズ (平均化)
→動画によって音量が異なるため
- 沈黙の位置で音声を区切る
→発言の整合性を損なわず、かつ長い文章になることを防ぐ
- 各音声を文字起こしし、それぞれの文章に感情分析

統合・補正

表情の分析で得た結果を、音声解析と自然言語処理で得た結果によって補正する。

例) 音声解析結果での補正

Angry:2.4%	Anger:40 → 1.40	Angry:3.36%	補正後
Happy:10.2%	Joy:5 → 1.05	Happy:10.71%	Angry:2.733%(3.33)
Neutral:74.8%	Calm:29 → 1.29	Neutral:92.492%	Happy:8.710%(1.49)
Sad:4.0%	Sorrow:26 → 1.26	Sad:5.04%	Neutral:75.225%(0.425)
Surprise:8.6%	Energy:32 → 1.32	Surprise:11.352%	Sad:4.099%(0.099)
		計 122.954%	Surprise:9.233%(0.633)
		それぞれを掛け合わせる	計 100%



プログラミング環境がないPCで使用することを想定し、Tkinterを用いて実行画面の可視化をした。

開発環境
Windows 11 Home
Python 3.8.8
Anaconda3 (64bit)

図4. 感情分析PCアプリ「Mia」サンプルイメージ

引用と参考文献

- [1]Pepper (ペッパー) | ロボット | ソフトバンク: <https://www.softbank.jp/robot/pepper/>
- [2]警察庁Webサイト: <https://www.npa.go.jp/publications/statistics/safetylife/jisatsu.html>
- [3]CNNの仕組み解説画像: <https://atmarkit.media.co.jp/ait/articles/1901/06/news032.html>
- [4]安倍内閣総理大臣記者会見 (令和2年2月29日): https://www.kantei.go.jp/jp/98_abe/statement/2020/0229kaiken.html
- [5]小林、原: ニューラルネットワークによる人の基本表情認識、計測自動制御学会論文集、Vol.29、No.1、112/118 (1993)
- [6]西路、遠藤、當間、山田、赤嶺: ディープニューラルネットワークによる画像からの表情表現の学習、The 29th Annual Conference of the Japanese Society for Artificial Intelligence (2015)