

音声認識を用いた理解の程度の評価

Estimation of the level of understanding using speech recognition

関根 航宇[†] 黒木 啓之[†]

Ko SEKINE[†] Takashi KUROKI[†]

[†] 東京都立産業技術高等専門学校

[†] Tokyo Metropolitan College of Industrial Technology

1. はじめに

実地での作業を指導する場合、口頭で指導した内容についての理解を確認することがある。このとき相手が気を遣って理解できていないのに理解できていると回答してしまう場合がある。しかし、指導している内容によっては、きちんと理解していないと危険な場合がある。そこで、相手が質問に回答した際の音声から相手が本当に理解しているかどうか確認できると便利である。

現在、感情に関する音声認識の分野では、平常時の音声と感情の込められた音声を比較する研究[1]や、感情を込めた発話の文頭の一句に含まれる韻律性を用いて感情を判断する研究[2]などが行われている。

そこで本研究では、質問の回答者が本当に理解しているかどうかを判定することができるようにすることを目的とする。

2. メル周波数ケプストラム係数

メル周波数ケプストラム係数(Mel-Frequency Cepstrum Coefficient,(以下,MFCC とする))は、音声特徴量抽出手法の1つである。メル尺度と呼ばれる音の高低に対する人間の感覚尺度をもとに周波数を変換し、そこからケプストラム分析によって特徴量を抽出するという手法である。

3. モデルアンサンブル

機械学習の多くは、与えられた課題に対して可能な限り正確な予測モデルを開発することに焦点を当てているが、1つのモデルを作成するのではなく、モデルの集合で構成される予測モデルをモデルアンサンブルと呼ぶ。

4. バギング

バギングは、アンサンブルを作成する標準的なアプローチの1つである。バギングを用いた場合アンサンブル内の各モデルは、データセットのランダム標本によって学習する。

5. ランダムフォレストモデル

ランダムフォレストは、機械学習の教師あり学習の一種である。バギング、特徴量集合の標本抽出、決定木を組み合わせたものである。ランダムフォレストモデルは個々のモデルが生成されると、アンサンブルは必要とされる予測の種類に応じて、多数決または中央値を返すことで予測を行う。

6. 研究方法

インターネット上には「怒り」「嫌悪」「恐怖」「喜び」「悲し

み」「驚き」の6つ感情を込められた音声のデータセットを無償で配布しているサイト[3]が存在する。それを機械学習モデルに学習させ人の音声から気分の上下などを判定する。以下に研究手順を示す。

1. インターネットに公開されている6つの感情の音声のデータセットからMFCCを用いて特徴量を抽出する。
2. 抽出した特徴量を機械学習モデルに学習させる。
3. 実際に音声データを入力して発話者の感情が6つのうちのどれに当てはまるかをモデルに判断させる。
4. モデルの判断の結果によって発話者が理解できているかどうかを判断する。

7. 研究結果

現在までに、ランダムフォレストモデルに、インターネット上で提供されている100人分の音声データを学習させ、ある人物の音声が入力された時に、その人物が100人の中で誰であるかを判断するということができています。

8. 今後の課題

今後の課題としては、感情が込められた音声データをランダムフォレストモデルに学習させ、音声データが入力されたときに発話者の感情を判断できるようにすることである。また、感情の判断結果から指導内容の理解の程度の判断を行う方法を模索する。

参考文献

- [1] 平賀裕, 斎藤善行, 森島繁生, 原島博, “音声に含まれる感情情報抽出の一検討,” 電子情報通信学会技術報告.HC, ヒューマンコミュニケーション, vol.90, no.16, pp.1-8, Apr. 1994.
- [2] 重永實, “感情の判別分析からみた感情音声の特性,” 電子情報通信学会論文誌 A, vol.J83-A, no.6, pp.726-735, Jun. 2000.
- [3] “辛徳泰, 江俊鋒, 高道慎之介, 斎藤佑樹, 相澤彰子, 猿渡洋, JNVN: a Japanese emotional speech corpus with both verbal content and nonverbal vocalizations (言語音声と非言語音声を持つ日本語感情音声コーパス), 高道慎之介, ”https://sites.google.com/site/shinnosuketakamichi/research-topics/jvnv_corpus, Oct. 2023.