

論文内容の要旨

博士論文題目 Comparison of Topic Classification Methods for Spoken Inquiries

(質問発話のトピックス分類手法に関する比較)

氏名 Torres Rodriguez, Rafael Antonio

(論文内容の要旨)

音声によるコミュニケーションは人にとって自然であり、音声認識技術は、自然なインタフェース(NUI)や対話システムとして、利用されている。とくに、スマートフォンの普及により、音声による質問に対する情報の提示が普及してきている。この論文では、音声による情報案内システム「たけまるくん」で収集された実データをもとに、音声認識技術によるトピックスの分類の機械学習アルゴリズムの研究を行っている。入力音声発話によるトピックスの同定は、適切な応答文や情報の提供にとって、重要な課題であり、この論文は、音声認識結果に基づく機械学習による多クラス分類への挑戦である。

実環境の音声認識は、話し言葉の認識の難しさによる認識誤りを多く含んでおり、さらに、入力も比較的短いものが多い。この論文では、「たけまるくん」の実データから頻度の多い15のトピックスをとりあげ、機械学習によるトピックス分類を行っている。機械学習の手法として、認識結果のBag of Wordsに基づくサポートベクターマシン(SVM)、最大エントロピー法(ME)、さらに、単語列情報を用いるPrefixSpan Boosting-Based法を取り上げている。いずれの手法も、従来の最近傍法よりも高いトピックス分類率である。機械学習の入力特徴として単語と文字単位のユニットを比較している。入力が比較的短いので、文字単位の入力が優れていることを示している。さらに、3つの手法の誤り傾向が異なるので、3つの手法を融合するスタック一般化法の利用を考案し、高い分類率を示している。

音声認識結果に基づくトピックス分類の性能を、音声認識誤りが少ない書き起こしを入力とした場合との比較し、また、比較的音声認識が高い大人の発声や、音声認識率の低い子供の発声に対する分類率の結果も示している。

氏名	Torres Rodriguez, Rafael Antonio
----	-------------------------------------

(論文審査結果の要旨)

音声によるコミュニケーションは人にとって自然であり、音声認識技術は、自然なインタフェース(NUI)や対話システムとして、利用されている。とくに、スマートフォンの普及により、音声による質問に対する情報の提示が普及し、入力音声発話のトピックスの同定は、重要な課題となっている。

この論文では、「たけまるくん」の実データから頻度の多い15のトピックスをとりあげ、音声認識結果に基づく3つの手法の機械学習でのトピックス分類を行っている。いずれの手法も、従来の最近傍法よりも高いトピックス分類率を示している。さらに、3つの手法の誤り傾向が異なるので、3つの手法を融合するスタック一般化法(Stacked Generalization)の利用を考案し、高い分類率を達成している。音声認識結果に基づくトピックス分類の性能を、音声認識誤りがなく書き起こしを入力とした場合と比較し、また、音声認識が高い大人の発声や、音声認識率の低い子供の発声に対する分類率の結果も示していることは、評価できる。これらの結果は、音声入力を利用した自然なインタフェース(NUI)や対話システムの作成にも大いに役立つものである。

これらの成果を、情報処理学会の学術論文、InterSpeechなどの重要な国際会議などで4編の論文として発表しており、国際的にも高い評価を得ている。また、日本語も読み書きおよび会話も日本人と同等であり、また、英語でのコミュニケーション能力もネイティブと同等です。

音声認識の分野を深く理解し、とくに、機械学習の各種アルゴリズムに精通した研究者として国際的な舞台での活躍が大いに期待できます。

平成24年12月18日に開催した公聴会の結果も参考にして、本博士論文の審査を行い、本論文は、博士論文(工学)として十分な価値があるものと判断した。