

論文 / 著書情報
Article / Book Information

題目(和文)	文脈を考慮した対話における感情認識
Title(English)	
著者(和文)	石渡太智
Author(English)	Taichi Ishiwatari
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12914号, 授与年月日:2024年9月20日, 学位の種別:課程博士, 審査員:徳永 健伸,岡崎 直観,村田 剛志,宮崎 純,齋藤 豪
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12914号, Conferred date:2024/9/20, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	情報工学 知能情報	系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of	(工学)
学生氏名： Student's Name	石渡 太智		審査員主査： Chief Examiner	徳永 健伸	

要旨（和文 2000 字程度）

Thesis Summary (approx.2000 Japanese Characters)

対話における感情認識は、会話における各発話の感情を認識する技術であり、ソーシャルメディアでの感情分析や感情的かつ共感的な対話システムの構築に利用される重要な技術として注目を集めている。対話の感情認識を実現するためには、話者が示す感情の種類を定義し、定めた感情ラベルを対話の各発話に付与して、作成したデータセットを基に対象発話の感情ラベルを予測する識別モデルを構築する必要がある。また、識別モデルを構築するためには、発話間の関係や会話の履歴、話者固有の特徴、常識的知識などの対話の感情認識タスク特有の課題に取り組む必要がある。

本研究は、対話の感情認識における上記の研究課題の中でも、特に識別モデルの構築に着目し、発話間の関係と会話の履歴を利用する新たな識別モデルを提案する。感情の種類と感情ラベルの付与に対して、本研究は、識別モデルの構築に取り組む従来研究と同様に、Ekman の 6 感情を基にしたベンチマークデータセットを用いる。識別モデルの構築では、課題の解決によって高い認識性能を示すことが期待され、多くの従来研究の注目を集める発話間の関係と会話の履歴に着目する。

まず発話間の関係の課題では、発話の距離を利用する手法を提案する。対話の感情認識では、発話間の関係の中で、自分自身の感情の推移を表す自己依存と他者の発話が影響を与える他者依存の関係が、話者の感情に影響を与えることが知られている。従来手法の多くは、グラフニューラルネットワークを用いて自己依存と他者依存の関係を利用し、高い認識性能を示した。しかしながら、これらの依存関係を利用する手法は、発話の距離を考慮しない課題が存在する。話者の感情は、しばしば直近の発話の影響を受けるなど、発話の距離に依存する。従って、発話の距離は感情認識を行う上で重要な情報である。そこで本研究は、発話間の関係に加えて、対象の発話から周辺の発話への距離の情報も利用する手法を提案する。提案手法を用いることで、自己依存と他者依存を含む発話間の関係と、発話の距離の両方を利用することができる。対話の感情認識における 3 つのベンチマークデータによる評価実験を通して、提案手法の有効性を確認した。

次に、会話の履歴の課題では、発話間の関係を利用する識別モデルと、会話の履歴を利用する識別モデルを組み合わせた手法を提案する。対話の感情認識では、同じ発話であっても、一連の会話の履歴に応じて異なる感情を示すことがある。会話の履歴を利用する代表的な方法として、連続した複数の発話を連結し、言語モデルに入力する方法がある。この手法は、識別対象の発話とその先行文脈に注意を向けるため、一連の会話の履歴を考慮することができる。しかしながら、この手法は、会話全体に注意を向けるため、逆に個々の発話の依存関係の利用が容易でない。

そこで、本研究は発話間の関係を利用するモデルと、会話の履歴を利用するモデルを組み合わせたアンサンブル手法を提案する。単純に組み合わせるだけでなく、過去の会話から会話の内容が近いものを検索し、動的な重み付き線形和によって補強する事例ベース手法を提案する。具体的には、識別対象の発話とその先行文脈をクエリーとして、会話の履歴の観点で意味的に近い発話を訓練データセットから k 近傍法を用いて検索する。

検索した発話（近傍事例）に付与された感情ラベルと、識別対象の発話との距離を基に感情ラベルの確率分布を作成し、発話間の関係を利用するモデルの確率分布と、動的な重み付き線形和によって組み合わせる。提案手法を用いることで、発話間の関係と会話の履歴の両方の特徴を利用することが可能となる。対話の感情認識の3つのベンチマークデータによる評価実験を通して、動的に重み係数を変更する提案手法が、最高水準の認識性能を示し、その有効性を確認した。

本研究の貢献を示す。本研究は、対話の感情認識において、発話と発話の距離の情報を利用するために、発話間の関係を利用するグラフニューラルネットワークに、はじめて距離の情報を付与する方法を提案した。提案手法を用いることで、自己依存と他者依存を含む発話間の関係と、発話の距離の両方の利用を可能にした。従来手法との比較実験を通して、提案手法は従来手法を上回る最高水準の認識性能を示し、その有効性を確認した。また、本研究は、発話間の関係を利用する識別モデルと、会話の履歴を利用する識別モデルを組み合わせる手法を提案した。2つの異なるモデルを組み合わせる方法として、近傍事例を活用し、はじめて対話の感情認識タスクに適用した。単純に組み合わせるだけでなく、識別対象の発話に応じて動的に変化する重み係数を用いて発話間の関係を利用するモデルの確率分布と、近傍事例による確率分布を組み合わせた。従来手法との比較実験を通して、重み係数を動的に変更する提案手法は従来手法を上回る最高水準の認識性能を示し、その有効性を確認した。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).

(博士課程)
Doctoral Program

論文要旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	情報工学 知能情報	系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of	(工学)
学生氏名： Student's Name	石渡 太智		審査員主査： Chief Examiner	徳永 健伸	

要旨（英文 300 語程度）

Thesis Summary (approx.300 English Words)

Interest in emotion recognition in conversations (ERC) has been increasing in various fields because it can be used to analyze conversations conducted via social media. In ERC, the relationships between the utterances and the conversational history are known to affect the speaker's emotions. In this research, we propose novel discriminative models considering the relationship between utterances and the history of the conversation.

Many recent ERC methods use graph neural networks (GNN) to consider the relationships between the utterances. In particular, the strong method considers self-speaker and inter-speaker dependencies in conversations by using relational graph attention networks. However, GNN does not consider the distance between utterances. In this research, we propose novel positional embeddings that provide GNN with positional information reflecting the graphical structures. Therefore, our proposed methods can capture both the speaker dependency and the distance. Experiments on ERC datasets show that our proposed method is beneficial to recognizing emotions expressed in conversations.

In ERC, some utterances with the same surface can show different emotions depending on the conversational history. A typical solution to this issue is to encode contextual information by concatenating a series of utterances and inputting them into language models. However, these methods do not consider the relationships between the utterances due to focusing attention on the entire conversation. In this research, we propose a method to incorporate the models considering utterance relationships to the models considering conversational history. Given a target utterance, we search the training dataset for utterances that are semantically similar to the target by using the models considering conversational history. The retrieved utterances are used to calculate probability distribution on emotion labels. The distribution is combined with another probability distribution from the utterance relationship model by using a weighted linear summation. Our experimental results on ERC datasets show that our method performs best over the baseline methods.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note: Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1 copy of 800 Words (English).

注意：論文要旨は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Tokyo Tech Research Repository Website (T2R2).