

ウェブニュース関連ツイートの極性に着目した 共感対話システムの提案と評価

田原 俊一[†]、池田 和史[†]、松本 一則[†]、帆足 啓一郎[†]

[†]KDDI 総合研究所 〒356-8502 埼玉県ふじみ野市大原2丁目1番15号

1. まえがき

近年、ユーザと雑談を行う対話システムが広く利用されている。ユーザの発話（ユーザ発話と定義）に対しシステムが共感を示す応答をすることは、ユーザの満足を向上する可能性がある。本稿ではこのような対話を実現するために、ニュースに関連するツイートをシステムの応答（システム応答と定義）として対話を行う手法を提案する。予備実験を通して、ユーザ発話と同じ極性(Positive, Negative)のツイートをシステム応答とすることで、ユーザの共感が高まることが示唆された。極性を考慮した提案対話手法と既存の対話手法の比較評価実験を行い、提案手法は既存手法と比較し、より共感度の高い対話が可能であることを確認した。

2. 関連研究

山口ら[1]は、学生とカウンセラーとの対話において、学生の発話とカウンセラーによる相槌を学習させ、発話が入力された際、適切な相槌を予測する手法を提案している。この手法は、システム応答が一定数であるため、応答の種類が乏しい。Jiwei ら[2]によれば、過度に学習した Deep Learning を用いた対話システムでは、「分からない」等の種類の数が乏しい応答をしがちになり、その結果対話が終了してしまうことが報告されており、良い応答ではないと主張している。従って、より共感度の高い対話システムを実現するためには、バリエーション豊かな応答が可能な手法が必要となる。

3. 提案手法

2章で述べた課題に対して、ニュースに関連するツイートをシステム応答として対話を行う手法を提案する。ユーザ発話に対してどのようなツイートを提示することが最も共感を生むかは自明ではないため、ユーザ発話の極性に対し、システムが同じ極性のツイートを応答すれば、ユーザの共感度は高まるという仮説を立てる。

仮説を検証するための対話システムの構成を図1に示す。次の手順でユーザとの対話を行う。(1)コンテンツ情報をニュース記事とし、ユーザの趣味嗜好

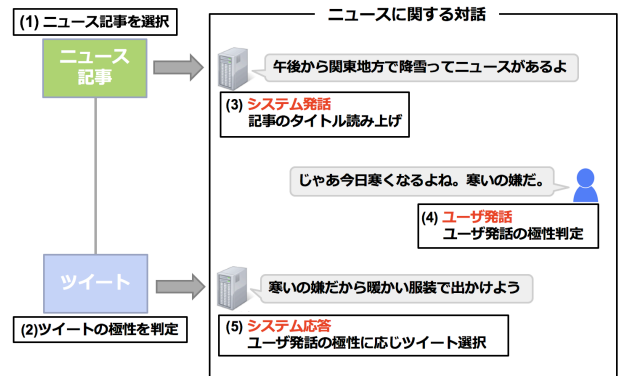


図1 提案手法における対話システムの構成

好に応じて記事を選択する。(2) 文の感情分析を行う Google API § を用いて記事に関連するツイートの極性を予め3極性 (Positive, Negative, Neutral) に分類する。(3) 対話開始時に記事のタイトルをシステムが読み上げる (システム発話)。(4) システム発話に対しユーザが任意の発話を行い、Google APIを用いて発話内容の極性の判定を行う。(5) ユーザ発話の極性を考慮して、(2)で3極性に分類したツイート群の中から1つを選択し、その群の中で特に主観表現を豊富に含むツイート1件を文献[3]の手法を用いて選択する。

4. 予備実験

4.1 実験データの収集と評価方法

3章で述べた仮説を検証する。例えば、ユーザの Negative な発話に対し、システムは Negative に応答する場合と Positive に応答する場合で、どちらがユーザに共感を与えるかを調べる必要がある。Yahoo!ニュースから取得した3件の記事と各記事に関連するツイート各45件を予備実験に利用する。ツイートは4名の作業員で極性を付与する。システム発話として記事のタイトルを、ユーザ発話及びシステム応答としてツイートをそれぞれ作業員とは別の被験者に提示する。ユーザ発話及びシステム応答それぞれに3極性を割り当て、計9条件を設定する。条件ごとに5対話を用意し、計45対話を実験データとする。

実験データの各対話に対し、「ユーザ発話に対するシステム応答がユーザに共感を感じさせたと思うか」という問い方で、被験者8名に対し、-3から3

Empathic Dialog System based on Sentiments Extracted from Tweets Posted for Web News

[†]Shunichi Tahara, Kazushi Ikeda, Kazunori Matsumoto, Keiichiro Hoashi, KDDI R&D Laboratories, Inc.

§ <http://cloud.google.com>.

までの共感度のアンケートに回答させる。表1に共感度と評価項目を記す。

4.2 評価結果

3件のニュース記事における、被験者8名による全対話の平均共感度の結果が表2である。ここでは、平均共感度が0よりも大きい場合、共感が得られたとみなすものとする。

ユーザ発話の極性が Positive, Negative の場合、それぞれ応答として Positive, Negative を示すとユーザの共感が得られることが分かった。また、ユーザ発話の極性が Neutral の場合、どの極性の応答をしてもユーザの共感が得られなかった。この結果から、提案手法ではユーザ発話が Positive または Negative であれば、同じ極性のツイートを応答し、ユーザ発話が Neutral であれば、文献[1]を参考に相槌を応答とする。

5. 提案手法の評価

5.1 評価方法

提案手法の有効性を評価するため、ユーザ発話に対しシステムが6種類の相槌の中から1つ応答する相槌手法と、ニュース記事に関連するツイートからランダムに1件選択し応答するランダム手法と比較し、評価実験を実施した。実験は対話システムのアプリケーションを搭載したスマートフォンで行う。また、システム発話→ユーザ発話→システム応答→ユーザ発話→システム応答を1対話とする。

ランダム手法と提案手法では、Yahoo!ニュースから取得した3件の記事と各記事に関連する引用ツイート計1033件を実験に利用する。

8名の被験者が実験に参加し、各被験者は相槌手法・ランダム手法・提案手法の順にそれぞれニュース記事1~3に関する対話を行う。各対話の実施後、「あなたの発話に対するシステムの応答に対して、共感を感じましたか?」という問い方で、表1と同じ評価項目のアンケートに回答させる。また各手法について自由記述による主観評価を求める。

5.2 評価結果

3件のニュース記事における、被験者8名による全対話の平均共感度の結果が表3である。表3では、平均共感度が0よりも大きい場合、共感が得られたとみなしている。

3手法の中で平均共感度が最も低い手法が相槌手法で、最も高い手法が提案手法であることが分かった。また、平均共感度が0を上回った手法は提案手法のみであることが分かった。

相槌手法に関しては「回答が面白くない」等の意見があり、ランダム手法に関しては「応答が一方向的に感じた」等の意見があった。

表1 共感度と評価項目

共感度	評価項目
-3	全く感じさせない
-2	ほとんど感じさせない
-1	どちらかというと感じさせない
0	どちらともいえない
1	どちらかというと感じさせる
2	ほとんど感じさせる
3	非常に感じさせる

表2 予備実験における平均共感度

		システム応答		
		Positive	Negative	Neutral
ユーザ発話	Positive	0.983	-1.217	-0.417
	Negative	-0.908	0.058	-0.583
	Neutral	-0.417	-0.583	-0.158

表3 各手法における平均共感度

	相槌手法	ランダム手法	提案手法
平均共感度	-0.125	0.000	1.000

提案手法に関しては「意味のある回答が多い」等の意見があった。提案手法による対話を実施した際のログを見ると、ユーザ発話に対して共感を示す応答が多かった。一方で、ユーザ発話とシステム応答の内容の極性は同じではあるが、それぞれ異なる内容であるため、文脈が合わずに共感度が低くなった対話例も見られた。

6. 結論と今後の課題

本稿では、ユーザ発話が Positive または Negative であれば、同じ極性のツイートを応答に利用し、ユーザ発話が Neutral であれば、相槌を応答とする手法を提案した。被験者実験によって、提案手法は既存手法と比較して、より共感度の高い会話が実現できることを確認した。今後の課題としては、違和感のない自然な対話を実現すべく、ユーザ発話に対して文脈を考慮した発話内容に関連性の高いツイートを選択し、応答する仕組みを構築する必要がある。

参考文献

- [1] 山口貴史, et al. "傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成." 人工知能学会論文誌 31.4 (2016): C-G31_1.
- [2] Jiwei Li, et al. "Deep reinforcement learning for dialogue generation." Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 1192-1202 (2016)
- [3] 池田和史, 服部元, 滝嶋康弘. "ニュースキュレーションサービスのためのネットコメント要約手法の提案." 第76回全国大会講演論文集 2014.1 (2014): 41-42.