

非テキスト情報の統合により円滑な対話を促進するチャットボットの設計と評価

2020SE013 磯谷和紀 2020SE064 高橋要太

指導教員：沢田篤史

1 はじめに

生成 AI の普及により AI によるチャットボットがビジネスや学習において盛んに活用されている。その例としてテキストベースのチャットボットによるメンタルヘルスケアでの活用事例などがある。

テキストベースのチャットボットにおいて、テキスト情報のみから応答生成を行うため、ユーザの感情に十分寄り添うことができていない問題がある。

本研究の目的は、テキスト情報と非テキスト情報を統合した感情予測によりチャットボットとの円滑な対話を実現することである。

この目的を達成するために本研究では、感情分析に影響を与える非テキスト情報を明らかにすることと、非テキスト情報を用いて会話を円滑にする機構を実現することの2つを技術課題に設定する。

上記の技術課題に対して、テキスト情報に打鍵速度を加えることでユーザの感情への親和性を高くすることと、打鍵速度とテキスト情報を統合した感情予測を行うチャットボットを提案することの2つのアプローチを取る。

本研究では提案した方法に対して、17名の被験者を対象に評価実験を行った。

本研究で提案したチャットボットは、怒りの文章に対しては低い評価になったが、悲しみと怒りの文章に関しては高い評価を得た。個性や満足度に関しても高い評価を得た。

2 チャットボットとその問題点

2.1 チャットボットの普及に伴う問題

生成 AI の普及により AI によるチャットボットが、学習や仕事をするうえで盛んに用いられている。その例として、カスタマーサービスや手続きの案内に活用されたり、メンタルヘルスケアとして活用されるなどして、チャットボットは私たちの生活を支えている。このような背景からチャットボットにはユーザの感情に寄り添うことが求められている。

チャットボットはシナリオ型と AI 型の2つに大きく分類される。シナリオ型は作成の段階で質問に対する応答を準備しておくことから、ある程度決まった質問に対する応答を行う。AI 型は質問に対して機械学習によって応答を生成するので、あらゆる質問への応答が可能である。2つのチャットボットでは応答可能な範囲に違いがある。これらテキストベースのチャットボットでは、ユーザの感情を正確に把握できないことが問題である。シナリオ型と AI

型ともにテキストベースのチャットボットはその大部分が、非テキスト情報が捨て去られているからである。

2.2 関連研究

Mohamed の研究 [1] では、感情分析を行うチャットボットと感情分析を行わないチャットボットの実装を行った。2つのチャットボットで Mohamed が設定した有用性・使いやすさなど 10 の評価項目を検討し2つのチャットボットの品質を比較することを目的とする。結果として、チャットボットの性格と人間性・ユーザ満足度が向上していることを実証している。

Zaijing ら [2] はテキスト、オーディオ、ビジュアルを特徴とし、感情ベクトルとして抽出する Decoder 部分を持たない Transformer ベースのモデル、Emofomer モデルを提案している。3つの感情ベクトルをそれぞれ文ベクトルと統合して EmoCaps という感情カプセルを作成し、最後にコンテキスト分析モデルによって話者の感情を決定するというものである。Zaijing らの提案したモデルを IEMOCAP データセットおよび MELD データセット上のほかの最先端モデルと比較を行った。結果として、提案モデルがほかのモデルに比べて高い F1 スコアとなり、高いパフォーマンスを発揮した。

感情分析を用いたチャットボットの研究として、小瀬木らは、マルチターン対話の中に現れる感情の変化をとらえ、相手の感情に合わせた予測をするモデルを提案している [3]。テキストだけでなく、音声や映像を組み合わせた場合での対話応答精度の比較を行っており、テキストだけよりも音声や映像を組み合わせた場合のほうが精度が高くなることを示している。感情の継続性など話者情報を対話応答に組み合わせた場合と組み合わせない場合で対話応答予測の精度比較も行っており、話者情報を組み合わせた場合のほうが精度が高いことを示している。結果として、感情情報を対話応答予測に反映させることは有効であると結論付けている。

以上の研究結果からテキスト以外に複数の情報を統合しチャットボットに用いることの有用性が示された。

非テキスト情報として打鍵速度からユーザの感情を推定する研究が行われている。渡部らの研究 [4] では、キーボード入力の際、入力者の感情の変化によってキーボードタッチの速さに変動があることを実証した。この研究では、平常時を想定する場合、被験者がテキストエディタを用いて思いついたことを入力させる。怒りを想定した場合は、過去に経験した不満や怒りを思い浮かべ、その事柄について入力させた。結果は、怒りを想定した場合のほうが、男女

ともに入力文字数が増えた。総合平均では平常時と比べて怒りを想定した場合のほうが約 1.2 倍入力文字数が増えている。また、感情が発話速度に与える影響を分析する研究も行われている。橋本 [5] は、人間の感情の中で、最も基本的な感情として区別しやすい、怒り、喜び、悲しみ、平常の 4 つの項目に感情を分け、発話速度と正負の感情から分類基準を設定し、実験した。結果は、怒りは発話速度の上昇し、負の感情を伴い、喜びは発話速度の上昇し、正の感情を伴い、悲しみは発話速度が低下し、負の感情を伴った。また、どの分類基準にも当てはまらないものは感情が特でない平常であった。

3 本研究の目的と技術課題

本研究の目的はテキスト情報と非テキスト情報を統合した感情予測によりチャットボットとの円滑な対話を促進することである。この目的を達成するため既存のチャットボットが抱えている非テキスト情報を十分に活用できていないという問題を解決する。

目的を達成するために本研究では、感情分析に影響を与える非テキスト情報を明らかにすること、非テキスト情報を用いて会話を円滑にする機構を実現することを技術課題に設定する。

感情分析を用いたチャットボットの研究では、小瀬木ら [3] のように音声・映像・テキストを組み合わせた複数の情報による感情分析が行われる。本研究では、感情分析に影響を与えるであろう非テキスト情報を明らかにし、非テキスト情報が感情分析結果にどのように影響を与えるのかを明らかにする。非テキスト情報を用いた感情予測によって円滑な対話を促進するチャットボットを実現する。

4 円滑な対話を促進するチャットボットの設計

テキストベースのチャットボットにおいてユーザの感情を読み取る情報として非テキスト情報が捨て去られている傾向がある。

本研究ではテキスト情報と非テキスト情報を統合して感情予測を行いチャットボットへ導入する。テキスト情報と非テキスト情報を統合した感情予測結果によって決まる挿入句を応答生成モデルへ受け渡す。非テキスト情報として打鍵速度を提案する。テキストを入力するのに要する 1 文字あたりの平均時間を打鍵速度と定義する。打鍵速度にはユーザの動きが反映されており、渡部らの研究 [4] から示されている通りユーザの感情を予測するうえで要素になりえると判断した。

渡部らの研究と橋本の研究 [5] から、打鍵速度をテキストの感情分析に組み合わせたユーザの感情分類は表 1 のようにまとめられる。橋本の研究での正の感情がポジティブ、負の感情がネガティブとする。その他はニュートラルとする。

表 1 ユーザの感情の多クラス分類

感情分析	打鍵速度	感情予測
ポジティブ	速い	喜び
ネガティブ	速い	怒り
ネガティブ	遅い	悲しみ
ネガティブ	その他	ニュートラル (ネガティブ)
ポジティブ	その他	ニュートラル (ポジティブ)
ニュートラル	その他	ニュートラル

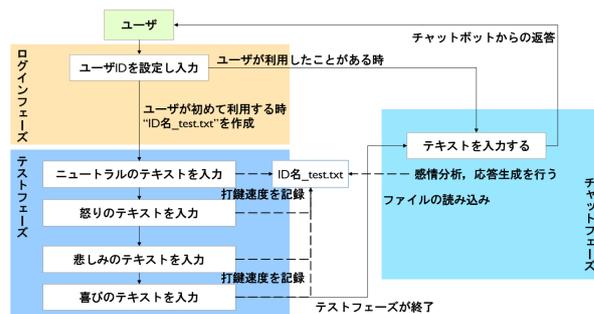


図 1 システム全体の設計

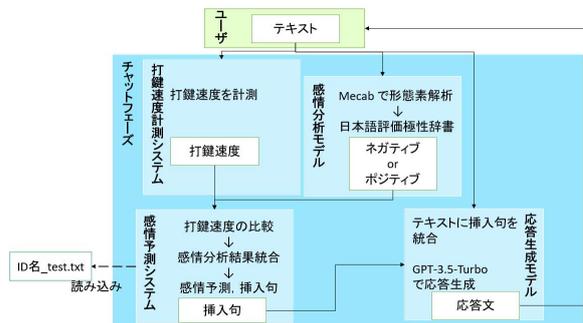


図 2 チャットフェーズの内部設計

4.1 システムの設計

本研究のシステム概要を図 1 に示す。提案システムは大きくテストフェーズとチャットフェーズの 2 つに分類できる。テストフェーズでは事前情報としてユーザの打鍵速度を記録する。チャットフェーズではテストフェーズで記録した打鍵速度を用いてチャットボットと会話を行う。テストフェーズで計測するユーザの打鍵速度は、ログインを行ったときに作成されるユーザファイルに記録する。テストフェーズでは、ユーザの事前情報としてニュートラル・怒り・悲しみ・喜びの時の打鍵速度を記録する。

チャットフェーズでは、初めに応答生成モデルに今から会話相手になってもらうこと、その際にユーザの感情予測

も送信することを伝えてから、チャットボットとの会話を始める。ファイルに記録している打鍵速度と、実際に会話を行うときの打鍵速度を比較しながらユーザの感情予測を行いチャットボットと会話を行う。チャットフェーズの概要を図 2 に示す。提案システムの役割は次のように分割できる。打鍵速度計測システムで打鍵速度を測る。感情分析モデルでテキストからネガティブ、ポジティブもしくはニュートラルの感情分析結果を返す。感情予測システムでテストフェーズ時の打鍵速度とチャットフェーズ時の打鍵速度の比較とネガティブ、ポジティブもしくはニュートラルの感情分析結果から応答生成モデルへ送るユーザ自身の感情予測と挿入句を決定する。応答生成モデルで挿入句とテキストから応答生成を行う。それぞれのシステム、モデルは独立したモジュールとして役割を果たすのでそれぞれを独立させた図 2 のような構造を提案する。

本研究では、感情分析モデルにテキストを単語ごとに分割し単語にスコアを付けてネガティブ、ポジティブもしくはニュートラルの感情分析結果を返す手法を用いた。単語にスコアを付けるために極性辞書を用いる。極性辞書とは単語や用語に対してネガティブやポジティブなどのラベル付けをした辞書のことである。提案するシステムでは、抜山ら [6] が用いていた日本語評価極性辞書 (用言編)、日本語評価極性辞書 (名詞編)*1 を用いる。以下この 2 つの辞書をまとめて、日本語評価極性辞書と表記する。

実装は抜山らの方法を参考にした。抜山らは janome を使って形態素解析を行っていたが、我々はチャットボット実装に技術を用いるためにより速く形態素解析を行える mecab を用いた。あらかじめテキストに mecab で形態素解析をかける。日本語評価極性辞書には単語が基本形で収録されているので、mecab で行うテキストの形態素解析も単語を基本形に変換する。その後は、日本語評価極性辞書内の単語と形態素解析後の単語のマッチングを行う。テキストは複数の単語で構成されており、テキスト中の単語 1 つが辞書のポジティブにマッチングした場合、単語の極性値が「1」としてテキスト全体の極性値に 1 足される。ネガティブにマッチングした場合は、単語の極性値が「-1」としてテキスト全体の極性値から 1 引かれる。ニュートラルにマッチングした場合はその単語の極性値を「0」とする。辞書に収録していない単語は「-」となり極性値は「0」とする。最終的な合計値でテキストから感情を決定する。一部、否定語がマッチングした場合は否定語の極性値を「0」にして、否定語の前後にある単語の極性値を反転する。

応答生成モデルでは、GPT-3.5-Turbo を用いて応答生成を行う。GPT-3.5-Turbo は GPT-3.5 系の中で一番能力が高くコストに対する効果が高いモデルである。*2 GPT-3.5 は、Brown ら [7] が開発した ChatGPT に使用されている自然言語モデルである。以前開発された GPT-2 に比べ

て、元の学習データセットが 40GB から 570GB に増加した。これにより、正確で洗練された結果を出力できるようになった。パラメータも GPT-2 の 15 億パラメータから、最大 1750 億パラメータを持つモデルが存在する。パラメータとは、機械学習モデルで学習データをもとに調整される重みやバイアスのことである。

4.2 チャットボットへユーザ自身の感情予測を導入

本研究では、ユーザ自身の感情予測を応答生成モデルに送ることでユーザの感情に寄り添った応答が生成されることを期待している。チャット内容とテキストの感情分析結果と打鍵速度を統合したユーザの感情予測結果によって決定する挿入句を応答生成モデルに送信し、応答生成モデルで応答を生成する。

テキストの感情分析結果と打鍵速度を統合したユーザの感情予測結果が喜びの場合は、「テキストからはポジティブ、文字の入力速度から感情分析結果は「喜び」と判定されました。喜びの感情に合わせた応答をしてください。」を挿入句としてユーザテキストに付け加えて応答生成モデルに送信し、それを踏まえた応答を生成する。

5 非テキスト情報を統合したチャットボットの実験と評価

本実験の打鍵速度と感情分析を統合して感情予測を行うチャットボットでは実験参加者の実験データを記録した。記録したデータはテキストと挿入句、打鍵速度、ユーザの感情予測、テキストからネガティブ、ポジティブもしくはニュートラルの結果、応答文である。ユーザのチャットデータの記録は考察を行うためのものである。

本実験は南山大学理工学部ソフトウェア工学科および理工学研究科ソフトウェア工学専攻に所属する、学生、大学院生の計 17 名を対象とした。被験者は打鍵速度と感情分析を統合したチャットボットとテキストから感情分析を行うチャットボット、感情分析を行わないチャットボットの 3 つのチャットボットと会話を行った。本研究の実験はテストフェーズ、チャットフェーズの 2 つのフェーズで構成される。テストフェーズで実験参加者の各感情における打鍵速度をデータとして記録する。チャットフェーズではテストフェーズで記録した打鍵速度を利用しチャットボットと各感情の話題で会話を行う。残り 2 つのチャットボットは、チャットフェーズのみ行う。3 つのチャットボットと会話を終えた後にアンケートを行う。

テキストから感情分析を行うチャットボットと感情分析を行わないチャットボットにおいてもログイン後、チャットフェーズに移行するとき、初めに応答生成モデルに、会話相手になってもらうことと情報を付け加える場合があることを送信し会話を始める。

実験の評価方法として Mohamed の研究 [1] で実験に使用された評価項目から「Personality and humanity(個性と人間性)」、「User satisfaction(満足度)」の 2 項目を抜粋

*1 東北大学の乾・鈴木研究室が公開している日本語評価極性辞書

*2 <https://platform.openai.com/docs/models/gpt-3-5>

し、ほかに各感情時の対処に関するアンケートなど評価項目を付け加えてアンケートを作成した。アンケートは、曖昧な応答を避けるために「とてもそう思う」、「そう思う」、「そう思わない」、「とてもそう思わない」の4段階で評価を行う。

評価するチャットボットは、打鍵速度と感情分析を統合したチャットボット、テキストから感情分析を行うチャットボット、感情分析をしないチャットボットの3つである。

本研究の実験結果としては、怒りの感情の時は全体的に適切な対処がされていなかったこと、悲しみと喜びの感情の時は打鍵速度と感情分析を統合したチャットボットと感情分析を用いたチャットボットが、適切な対処をされていたと答えられた。チャットボットの個性と満足度においても総合的に打鍵速度と感情分析を統合したチャットボットと感情分析を用いたチャットボットの評価が高く、2つを比べた場合、わずかに打鍵速度と感情分析を統合したチャットボットのほうが評価が高かった。

6 考察

実験時ユーザの感情予測の結果が期待するものと異なった原因に、打鍵速度の記録が期待と異なること、テキストの感情分析結果が期待と異なることが挙げられる。ニュートラルの打鍵速度を記録するときに打鍵速度に差がない、もしくは「怒り」時の打鍵速度が「ニュートラル」時の打鍵速度より遅いなど期待と異なる記録をしたときに、打鍵速度による感情の分類が想定と異なるので、チャット時の打鍵速度にも影響され感情予測にも影響が与えられた。チャット時の打鍵速度の速い、遅いの分類が期待するものであった場合でも、テキストの感情分析結果が期待と異なることでユーザの最終的に感情予測にも影響が与えられ期待と異なった。一方で、テストフェーズ時の打鍵速度とチャットフェーズ時の打鍵速度で記録が大幅に異なる被験者もいた。このことから、1回の記録で各感情時の打鍵速度を決定しないで、打鍵速度を更新していくことは有効に機能する可能性がある。打鍵速度の記録が期待されるものかつ、テキストの感情分析が期待するものに限っては11個のテキスト中7個が期待する感情予測に分類されたので、打鍵速度の記録方法の変更、感情分析モデルの変更、辞書の更新を行うことで期待する感情予測結果の数は多くなる可能性がある。今後の課題として、打鍵速度をオンライン学習で記録して打鍵速度の平均を更新していくことや、極性辞書の更新を行い感情分析の精度を高めること、感情分析モデルをTransformerなど文脈を理解できるモデルに変更し打鍵速度を統合する方法を用いることを挙げる。

7 おわりに

生成AIの普及によりAIによるチャットボットが、ビジネスや学習において盛んに活用されている。その例として、カスタマーサービスや手続きの案内、メンタルヘルスケアとして活用されるなどして、チャットボットは私たち

の生活に関わっている。このような背景からチャットボットには人間の感情に寄り添うことが求められている。一方で、テキストベースのチャットボットには、まだユーザの感情に十分に寄り添うことができていない問題がある。

本研究の目的は、テキスト情報と非テキスト情報を統合した感情分析によりチャットボットとの円滑な対話を実現することである。本研究の目的を達成することにより、チャットボットと会話が従来のもよりもユーザの感情に寄り添え、PCやスマートフォンなどテキスト入力を用いたチャットボットに活用が見込める。

この目的を達成するために本研究では、打鍵速度とテキスト情報を統合した感情分析を行うチャットボットを実装した。チャットボットの評価を行うためにテキストから感情分析を行うチャットボットと感情分析を行わないチャットボットを実装し被験者実験を行った。

本研究の結果として、打鍵速度と感情分析を統合したチャットボットは、会話を円滑にする有効性が見られた。怒りの文章には、やや低めの評価となり、悲しみと喜びは高い評価を得た。チャットボットの個性と満足度に関しても高い評価を得た。

今後の課題として、打鍵速度をオンライン学習で記録して打鍵速度の平均を更新していくことや、極性辞書の更新を行い感情分析の精度を高めること、感情分析モデルにTransformerなど文脈を理解できるモデルに変更し打鍵速度を統合する方法を用いることを挙げる。

参考文献

- [1] Mohamed Alaa Eddine Cherni: "Supporting chatbot intent recognition with sentiment analysis", *Master of Science in Computer Science at the Ruhr West University of Applied Sciences*, 2022.
- [2] LI, Zaijing, et al. "Emocaps: Emotion capsule based model for conversational emotion recognition." *arXiv preprint arXiv:2203.13504*, 2022.
- [3] 小瀬木 悠佳, 立石 修平, 大杉 康仁, 狩野 悌久, 中辻 真: "マルチターン対話における感情予測を活用した学習モデルの提案", 人工知能学会全国大会論文集第37回, 2023.
- [4] 渡部 容子, 小野田 誠, 能地 保: "キーボードタッチ入力による感情情報収集システムの男女差の検証", 第68回全国大会講演論文集, pp. 107-108, 2006.
- [5] 橋本 和希: "発話速度を考慮した自然発話の感情分析", naistar, 2008.
- [6] 抜山 雄一, 喜多 敏博: "オンライン授業におけるディスカッションの理解を促進するテキストマイニング手法の検討と実装", 日本教育工学会研究報告集, 2022.
- [7] BROWN, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems*, 2020, 33: 1877-1901.