

# ChatGPTを用いたメッセージ修正案の提示による対話相手からの好感度改善

## Improvement of Favorability from Dialogue Partners by Presenting Message Modification Suggestions Using ChatGPT

野呂 悠斗<sup>1\*</sup> 砂山 渡<sup>2</sup> 服部 峻<sup>2</sup>  
Yuto Noro<sup>1</sup> Wataru Sunayama<sup>2</sup> Shun Hattori<sup>2</sup>

<sup>1</sup> 滋賀県立大学大学院工学研究科

<sup>1</sup> Graduate School of Engineering, The University of Shiga Prefecture

<sup>2</sup> 滋賀県立大学工学部

<sup>2</sup> School of Engineering, The University of Shiga Prefecture

**Abstract:** With the increase in the number of SNS users, messages and posts that hurt or offend others have become a problem. These messages are not limited to general messages such as abusive language and slanderous remarks, but also include messages that are offensive only to specific people. Although there has been research on acquiring information on the speaker's preferences, there has been little research on generating message modification plans. In this study, we aim to realize better communication by creating a system that uses ChatGPT to extract the preferences of the interlocutor from the dialogue history and propose a sentence modification plan tailored to the interlocutor.

## 1 はじめに

近年、ソーシャルネットワーキングサービス（SNS）等のサービスの利用者が増加傾向にある。総務省による調査 [1] では、日本のソーシャルメディア利用者数は 2028 年には 1 億 1360 万人にのぼると予測されている。

SNS の利用におけるデメリットとして、他者を傷つける、あるいは不快にさせるメッセージが書き込まれる可能性が挙げられる。暴言や誹謗中傷は大多数の人にとって不快と感じられるものであり、客観的に適切でない場合が多い。しかし、一般的には問題がない場合でも、対話相手にとっては傷つく話題、嫌な話題であることがあり、メッセージ単体のみから客観的に見て不適切であると判定することが難しい場合がある。

そこで本研究では、二者のテキスト形式の対話環境での利用を想定した、ChatGPT を用いた対話相手の好感度を高められるメッセージ修正案の提示システムを提案する。ユーザと対話相手の対話履歴をもとに、ユーザのメッセージに対する対話相手の反応を分析し、対話相手が好む話題や嫌う話題、関心のない話題の嗜好情報を生成する。その後、新しく送信するメッセージ

に対し、対話相手への興味、関心、好意を表す客観的基準および嗜好情報をもとにした主観的基準の 2 つの観点から好感度の評価を行う。必要であれば、メッセージの修正案をユーザに提示する。修正案をもとにユーザはメッセージを修正し、それを対話相手に送信することで、対話相手のユーザに対する好感度を高める。また、本システムを利用することによる対話相手の好感度の向上について検証する。

## 2 関連研究

パーソナライズされた嗜好プロフィールに関する研究 [2] や、嗜好プロフィールに対してユーザが不快に感じるコンテンツをフィルタリングする研究 [3] がある。これらの研究では、インターネット上のプラットフォームでの閲覧履歴から、ユーザの嗜好を作成するシステムや、ユーザを不快にするレコメンドを嗜好プロフィールからフィルタリングするシステムが提案されているが、人間同士のコミュニケーションには対応していない。本研究では、対話相手の好感度を高めるための手法として、入力メッセージの修正案の提示を行う。

発話に含まれる品詞の関係から、名詞に対する感情を推定する研究 [4] や、発話の構文情報や音響的情報を

\*連絡先：滋賀県立大学大学院工学研究科電子システム工学専攻  
〒 522-8533 滋賀県彦根市八坂町 2500  
E-mail: on23ynoro@ec.usp.ac.jp

表 1: 主観的基準要素の説明

種類	説明
ポジティブ	対話相手に関心を示し、かつ好む話題
ネガティブ	対話相手に関心を示すが、好まない話題
ニュートラル	対話相手に関心を示し、かつ「ポジティブ」にも「ネガティブ」にも当てはまらない話題
無関心	対話相手に関心を示さない話題

もとに発話対象の好感度を計算する研究 [5], 単語の好感度と文法構造から感情を解析, 生成する研究 [6] がある. これらの研究では, 構文情報をもとに好感度や感情の推定を行っているが, 文章全体の内容を把握することは難しいことや, 好感度の計算の精度には限界があり, 誤った感情推定が行われることがある. 本研究では, より高精度な好感度推定を行うために ChatGPT を用いて対話相手の嗜好情報を作成し, 好感度を判定する.

### 3 本研究における好感度の定義

#### 3.1 客観的基準による好感度

本研究における客観的基準として, 対話相手に依存しない共通の指標を用いる. Horsham らの研究 [7] によると, 他者への好奇心と好感度には正の相関があることが示されている. また, 他者に対する好感度の変化と相手への理解度の関係に関する研究 [8] によると, 対話相手が自身を理解していると認識することで好感度が高まるとされている. 本研究では, 「共感」, 「関心」, 「好意」, 「ポジティブ」, 「親しみやすさ」の5つの要素について, 対話相手に依存しない内容を基準とした好感度を, 客観的基準による好感度と定義する. 「関心」については, 内容に対して否定的である場合を想定し, 「ポジティブ」の要素を加える. また, SNSでの利用を想定し, 親しみやすい文章である必要があるため, 「親しみやすさ」の要素を加える.

#### 3.2 主観的基準による好感度

本研究における主観的基準として, ユーザが入力したメッセージに対して対話相手がどう感じるかを判定する上で, 対話相手の特徴による指標を用いる. 本研究では, 「ポジティブ」「ネガティブ」「ニュートラル」「無関心」の4つの要素について, 対話相手に依存する内容を基準として用意し, この基準による好感度を, 主

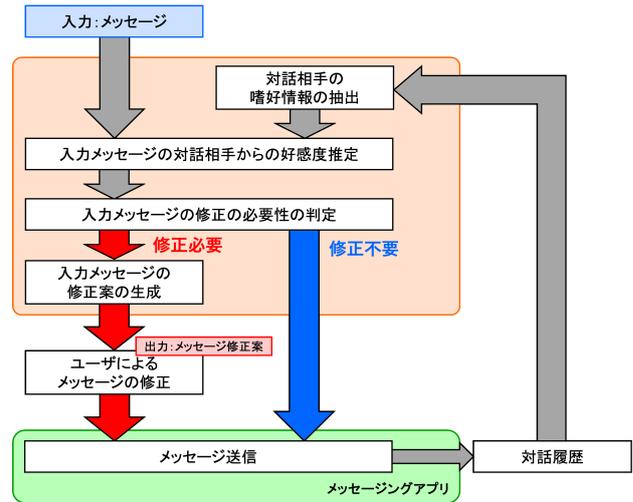


図 1: 対話相手の好感度を高められるメッセージ修正案提示システムの構成

観的基準による好感度と定義する. 表 1 に, 各主観的基準の種類の特徴を示す. ユーザが入力したメッセージに対して対話相手に関心を持つか, および好むかどうかにより4つに分ける. 「ネガティブ」と「無関心」の違いについて, 「ネガティブ」は話したくない話題, 「無関心」は話したくないわけではなく, そもそも関心がない話題であるという違いがある.

### 4 対話相手の好感度を高められるメッセージ修正案提示システム

本システムは, 「LINE」等の1対1での対話を想定している. ユーザの入力するメッセージが相手を傷つける, あるいは不快にさせるかどうかを2つの観点から ChatGPT により判定し, それをもとに修正案を提示する.

#### 4.1 対話相手の好感度を高められるメッセージ修正案提示システム

図 1 に本システムの構成を示す. 「LINE」等の1対1でのテキストによるメッセージのやり取りを想定する. ユーザが送信しようとしているメッセージと, ユーザと対話相手との対話履歴を入力データとしてシステムに入力する. 対話相手の嗜好情報の抽出では, 対話履歴のユーザのメッセージのグループ分けをし, それぞれの特徴を嗜好情報として抽出する. 入力メッセージの対話相手からの好感度推定では, 客観的基準と, 嗜好情報からなる主観的基準の2つの基準から, ユーザ

表 2: 対話相手のメッセージから関心・無関心を判定するプロンプト

```
文章から心理分析を行うプロとして、Aさんの「[messageA]」という発言に対し、Bさんの「[messageB]」という発言が、相手に関心のあるものならば「1」を、無関心ならば「0」を出力してください。アルファベットと文章はコマンドで区切ってください。
```

表 3: 対話相手のメッセージから関心・無関心を判定するプロンプト

```
文章から心理分析を行うプロとして、Aさんの「[messageA]」という発言に対し、Bさんの「[messageB]」という発言が、相手に関心のあるものならば「1」を、無関心ならば「0」を出力してください。アルファベットと文章はコマンドで区切ってください。
```

が入力したメッセージの好感度を推定する。そしてその結果をもとに、入力メッセージの修正が必要かどうかを判定し、修正が必要と判定された場合、メッセージの修正案の生成に進み、修正案を作成し、ユーザに提示する。修正が不要と判定された場合、入力メッセージをそのまま対話相手に送信する。ユーザによるメッセージの修正では、提示された修正案をもとにユーザ自身で修正前のメッセージを修正し、対話相手に送信することができる。

## 4.2 対話相手の嗜好情報の抽出

### 4.2.1 対話履歴のメッセージのポジネガ分類

ポジネガ分類を行う上で、入力としてユーザと対話相手との対話履歴（テキスト形式）を用いる。対話相手の反応から、ユーザのメッセージが対話相手にとって好む話題か、嫌う話題か、無関心な話題かを判定することで、ユーザのメッセージを「ポジティブ」、「ネガティブ」、「ニュートラル」、「無関心」の4つのグループに分類する。ここで分類を行うのは、分類後の各グループのメッセージ集合について、4.2.2 項で嗜好情報を作成するためである。

まず、ユーザと対話相手との全対話履歴からユーザのメッセージと、それに対する対話相手の返信メッセージを1組とするデータを作成する。1組のメッセージにおいて、ユーザのメッセージが対話相手にとって関心があるかどうかを、対話相手のメッセージをもとにChatGPTにより判定する。対話相手にとって無関心であると判定された場合、ユーザのメッセージは「無関心」のグループに分類される。一方、関心があると判定された場合、次のポジネガの判定に移る。

対話相手にとって関心があると判定されたユーザのメッセージは、対話相手が好む「ポジティブ」なメッセージであるか、嫌う「ネガティブ」なメッセージであるか、そのどちらでもない「ニュートラル」なメッセージであるかを対話相手のメッセージをもとにChatGPT

表 4: 対話相手のメッセージからポジネガを判定するプロンプト

```
#指示文  
文章から心理分析を行うプロとして、Aさんの「[messageA]」という発言に対し、Bさんの「[messageB]」という発言が、ポジティブな表現であれば「p」、ネガティブな表現であれば「n」を、「p」にも「n」どちらにも当てはまらなければ「e」を出力してください。アルファベット1文字のみ出力し、それ以外の文章は削除してください。
```

表 5: 嗜好情報を作成するプロンプト

```
#指示文  
あなたは文章から心理分析を行うプロです。文章を分析し、共通する話題や特徴を特定し、箇条書きで出力してください。
```

により判定し、判定結果をもとに、ユーザのメッセージを「ポジティブ」、「ネガティブ」、「ニュートラル」のいずれかのグループに分類する。

これらの判定結果をもとに、対話相手の反応からユーザのメッセージを「ポジティブ」、「ネガティブ」、「ニュートラル」、「無関心」の4つのグループに分ける。

表3に、対話相手のメッセージから関心・無関心を判定するプロンプトを示す。「messageA」にユーザのメッセージを、「messageB」に対話相手のメッセージを入れ、プロンプトを入力してChatGPTに送信することで「1」か「0」が出力される。出力が「0」の場合、ユーザのメッセージは対話相手にとって「無関心」とあると判定し、「1」の場合、関心があると判定する。例となるメッセージとして、「関心がある」例として、相手のメッセージに反応を示すメッセージを、「無関心」の例として、相手のメッセージの内容に対して無関心なメッセージを設定している。

表4に、対話相手のメッセージからポジネガを判定するプロンプトを示す。「messageA」にユーザのメッセージを、「messageB」に対話相手のメッセージを入れ、プロンプトを入力してChatGPTに送信することで「p」、「n」、「e」のいずれかが出力される。出力が「p」の場合、ユーザのメッセージは対話相手にとって「ポジティブ」とあると判定する。出力が「n」の場合、ユーザのメッセージは対話相手にとって「ネガティブ」とあると判定する。出力が「e」の場合、ユーザのメッセージは対話相手にとって「ニュートラル」とあると判定する。例となるメッセージとして、「ポジティブ」の例として、相手のメッセージに肯定や共感、興味を示すメッセージを、「ネガティブ」の例として、相手のメッセージの内容に対して否定的な内容や落ち込む内容のメッセージを設定している。

表 6: 客観的基準の「共感」を判定するプロンプト

文章から心理分析を行うプロとして、入力された文章について、相手の発言に共感や理解を示す程度を 5 点満点で評価してください。1 桁の数字で出力してください。

<5 点>  
 ...  
 <4 点>  
 ...

表 7: プロンプトで使用する客観的基準 5 項目の説明文

項目	説明
共感	共感や理解を示す程度
関心	関心・興味を持つ程度
好意	好ましく思う程度
ポジティブ	積極的・肯定的な度合い
親しみやすさ	親しみを感じさせる度合い

#### 4.2.2 嗜好情報の作成

4.2.1 項でグループ分けされたユーザのメッセージを入力とし、各グループ内のユーザのメッセージに共通する話題や特徴を抽出する。これにより、「ポジティブ」、「ネガティブ」、「ニュートラル」、「無関心」の 4 つの対話相手の嗜好情報が作成される。

表 5 に、特徴を作成するプロンプトを示す。各グループのメッセージの集合をプロンプトの #参照 に入力し、ChatGPT に送信することで、各グループの特徴を箇条書きで出力させる。プロンプトでは抽象的な表現にする指示を与えることにより、重要度の高い情報のみの特徴への反映を図っている。

### 4.3 入力メッセージの対話相手からの好感度推定

対話相手に送るために入力されたメッセージについて、対話相手からの好感度を、前節で用意した嗜好情報を用いた主観的基準、ならびに客観的基準を用いて推定する。

#### 4.3.1 客観的基準による好感度推定

本項では、ユーザが対話相手に送るために入力されたメッセージについて、対話相手からの好感度を客観的基準を用いて推定する方法について述べる。

新しく入力されたメッセージについて、3.1 節で定義した 5 つの客観的基準の項目について、その項目を満たす度合いを 1-5 の 5 段階で評価する。5 が最も良く、1 が最も悪い評価となる。本システムでは、客観的基準の 5 項目それぞれについて、点数が「4」以上であればその項目を満たすものとし、「3」以下であれば満たさないものとする。その後、5 項目のそれぞれの点数を出力する。

表 8: 抽出した特徴による入力メッセージの関心・無関心、およびポジネガを判定する判定プロンプト

#指示文  
 文章から心理分析を行うプロとして、「[入力メッセージ]」という発言が、相手に関心がありポジティブな表現ならば「p」を、相手に関心がありネガティブな表現ならば「n」を、相手に関心があり中立な表現ならば「e」を、無関心ならば「a」を出力してください。{#参照} を例にして出力してください。

#参照  
 p(ポジティブ)  
 ...  
 n(ネガティブ)  
 ...  
 e(中立)  
 ...  
 a(無関心)  
 ...

表 6 に、客観的基準の「共感」を判定するプロンプトを示す。新しく入力されたメッセージをプロンプトに入力し、ChatGPT に送信することで、判定結果に応じて 1-5 の数値を出力させることができる。各点数について基準となる文章を例文として設定することにより、点数のコントロールを行っている。なお、表 6 で示したのは「共感」を判定するプロンプトであり、他項目を判定する際は、表 7 に示す説明文をプロンプトの該当箇所に入力し、各点数の例を項目に合わせて変更する。

#### 4.3.2 主観的基準による好感度推定

4.2.2 項で得られた特徴をもとに、入力されたメッセージの内容が「ポジティブ」、「ネガティブ」、「ニュートラル」、「無関心」のいずれに当てはまるかを判定する。対話相手に送信しようとしたメッセージと、4.2 節で抽出した対話相手の嗜好情報を入力とし、メッセージが対話相手の「ポジティブ」、「ネガティブ」、「ニュートラル」、「無関心」の嗜好情報のうち、どれに当てはまるかを判定する。

表 8 に、抽出した特徴によるメッセージ判定のプロンプトを示す。新たに入力されたメッセージと 4.2.2 項で得られた特徴をプロンプトに入力し ChatGPT に送信することで、「p」、「n」、「e」、「a」のいずれかを出力させる。そしてその出力をもとに「ポジティブ」、「ネガティブ」、「ニュートラル」、「無関心」を判定する。

### 4.4 入力メッセージの修正の必要性の判定

4.3 節で得られた入力メッセージに対する好感度の結果をもとに、入力されたユーザのメッセージについて、以下の 3 つの条件について判定を行い、修正の必要の有無を決定する。条件 1 から条件 3 のうちいずれか 1

表 9: 入力メッセージを修正するプロンプト

	プロンプト
基本文	あなたは文章から心理分析を行うプロです。入力された文章について、次の条件に従って修正した文章を1つ作成してください。 - 入力された文章と同程度の文字数にしてください。 - 相手に対して親しみを感じさせる文章にしてください
条件 1	- 相手の発言に共感や理解を示す文章にしてください
条件 2	- 相手に対して好ましく思う文章にしてください - 相手に対して積極的・肯定的な文章にしてください - 相手の発言に相手の発言に対して関心・興味を持つ文章にしてください
条件 3	- {#参照}を参考にし、ポジティブな表現を増やした文章にしてください - {#参照}を参考にし、喜ばない表現、無関心な表現を避けた文章にしてください

つでも満たした場合、入力メッセージの修正が必要であると判定する。

- 条件 1 客観的基準の「親しみやすさ」を満たさない
- 条件 2 客観的基準の「共感」を満たさない、かつ「関心」と「ポジティブ」のどちらかを満たさない、かつ「好意」を満たさない
- 条件 3 主観的基準（嗜好情報）の「ネガティブ」または「無関心」に当てはまる

条件 1, 条件 2 については 4.3.1 項の結果をもとにしており、条件 3 については 4.3.2 項の結果をもとにしている。

条件 1 を設けた理由は、SNS 等での運用を想定しており、硬い文章ではなく、軟らかい表現の使用を促すためである。条件 2 を設けた理由は、相手に対して興味を持つような表現を増やすことにより対話相手を感じる好感度を高めるためのものである。条件 3 を設けた理由は、対話相手が嫌う、あるいは無関心な話題を話すことにより対話相手の好感度を下げることが避けるためである。

#### 4.5 入力メッセージの修正案の生成

4.4 節の修正判定結果をもとに、文章の修正に用いるプロンプトを改訂し、文章の修正を実施する。本システムでは、ChatGPT にプロンプトを送信し、修正案を1つずつ、合計3つの修正案を生成する。表9にメッセージ修正に用いるプロンプトを示す。基本文の内容に加え、条件 1 から条件 3 の修正条件のうち、満たしているものに対応する文章をプロンプトに追加する形である。条件 2 を満たしている場合には、条件 1 と条件 3 のうち満たしている修正条件に加え、以下の修正案を行う。1つ目の修正案は「共感」のみ、2つ目の修正案は「関心」と「ポジティブ」のみ、3つ目の修正案

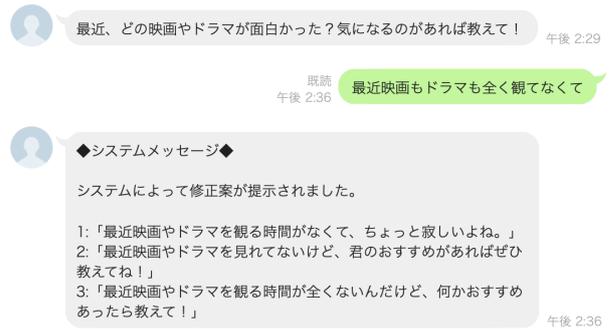


図 2: LINE 上での対話相手の好感度を高められるコミュニケーション支援システムのデモ

は「好意」のみを修正させるようにプロンプトに文章を追加する。

4.5 節で作成された3つの修正案がユーザに提示される。ユーザは提示された修正案をもとに、必要に応じてメッセージを追加で修正する。その後、メッセージを対話相手に送信する。図2に、LINE 上での対話相手の好感度を高められるコミュニケーション支援システムのデモの様子を示す。修正が必要な場合、システムはユーザに対して修正案を提示する。

図2に、LINE 上での対話相手の好感度を高められるコミュニケーション支援システムのデモの様子を示す。修正が必要であった場合、作成された修正案がユーザに提示される。ユーザは提示された修正案をもとに、必要に応じてメッセージを追加で修正する。

#### 4.6 ユーザによるメッセージの修正

4.4 節で提示された修正案をもとに、ユーザは新しく送信するメッセージを作成することができる。

修正案がそのまま利用できるとユーザが判断した場合、ユーザがそのメッセージをそのまま送信することができる。

修正案をそのまま利用できないが、一部修正すれば利用できるとユーザが判断した場合、ユーザは提示された修正案を修正して送信することができる。

ユーザが送信したくないと感じる、あるいは送信したい内容がないなどの理由で、修正案を利用しないとユーザが判断した場合、ユーザは元々送信しようとしたメッセージをそのまま送信したり、修正案を参考にせず元のメッセージを修正して送信することができる。

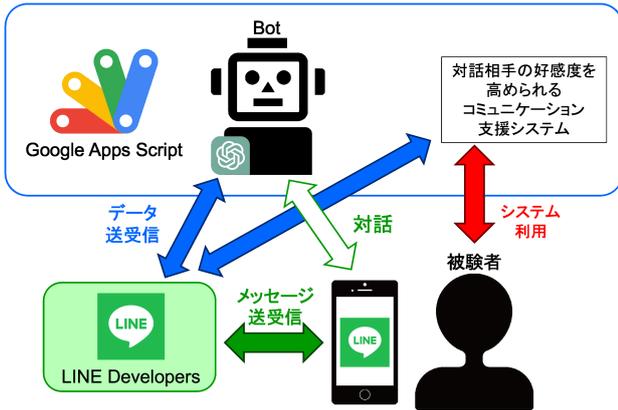


図 3: メッセージ修正案の評価および修正案に対する対話相手の嗜好情報の反映度の評価で用いる環境の構成図

表 10: 「太郎さん」のメッセージ生成のプロンプト

```
#命令文
あなたは 20 代の友人を相手にした対話をします。20 代の人間になりきって、返事(セリフ)を 1 つだけ作ってください。{#制約条件}に従ってチャットボットとして振る舞ってください。

#制約条件
- 短めの文章で、感情を込めて話してください。
- 話題が尽きそうなどときには、自分で新しい話題を出してください。
- 50 文字以下にしてください。
- カウントした文字数は出力に含めないでください。
- メッセージの内容については、{#特徴}に従ってください。

#特徴
- 相手のメッセージの大学生活やスポーツ観戦に関する話題には興味を示してください。
- 相手のメッセージのゲームや遊びに関する話題には好きではないことを示してください。
- 相手のメッセージのアニメやドラマに関する話題には興味を示さないでください。
```

## 5 メッセージ修正案の提示による対話相手からの好感度改善の検証

本章では、提案システムについて、被験者のメッセージに対してシステムを適用することにより、対話相手の好感度を高められたかどうかについての検証を行う。

### 5.1 目的

被験者とチャットボットの一对一の対話で、被験者のメッセージに対して提案システムを利用した場合と利用しない場合での比較を行う。対話相手の対話相手に対する好感度の違いに着目し、好感度の違いや嗜好情報が修正案に与える影響について検証する。

図 3 に、実験 2 で使用する環境の構成図を示す。被験者と対話するチャットボットおよび提案システムは Google Apps Script で実装しており、LINE Developers

表 11: 「太郎さん」の嗜好情報

グループ	特徴
ポジティブ	- 大学生活について - スポーツ観戦について
ネガティブ	- ゲームについて - 遊びについて
ニュートラル	- 気分転換の方法について - 身体の健康について
無関心	- アニメについて - ドラマについて

側のリクエストにより呼び出され、Bot の返信メッセージおよび提案システムによる修正案を LINEbot に返信させる。

### 5.2 対話相手

本実験では、「太郎さん」という名前の対話相手のチャットボットを準備した。表 10 に「太郎さん」のメッセージを生成するプロンプトを、表 11 にあらかじめ設定した「太郎さん」の嗜好情報を示す。プロンプトには特定の話題に対して興味を示す、好まない、興味がない設定を追加しており、「太郎さん」の嗜好情報に合わせたものとなっている。

メッセージ生成の際に用いる対話履歴は直近 3 往復のメッセージであり、それぞれ user と assistant に入力する。実験では、「太郎さん (A)」という名前のチャットボットと「太郎さん (B)」という名前のチャットボットを用意しているが、「太郎さん」のメッセージを生成するプロンプトに違いはない。「太郎さん (A)」との対話では、被験者が送信するメッセージに対して提案システムにより修正の必要性の有無の判定、修正案の提示を行う。一方、「太郎さん (B)」ではシステムの利用は行わず、修正案が提示されることはない。

### 5.3 実験手順

本項では、実験手順について述べる。実験の被験者は 15 名の大学生と大学院生であり。本実験で使用した ChatGPT のモデルは、すべて「gpt-4o-2024-08-06」である。

まず、被験者 15 名をグループ 1 とグループ 2 の 2 つに分ける。その後、グループ 1 の被験者は「太郎さん (A)」、「太郎さん (B)」の順、グループ 2 の被験者は「太郎さん (B)」、「太郎さん (A)」の順番で対話を行う。対話は最初にユーザが「こんにちは」というメッセージを送信するところからスタートし、その後は自由に対話を行ってもらった。各対話は 20 往復に到達すると終了する。各対話相手について、対話相手が送信する文章の修正前後での客観的基準および主観的基準の変化を調べた。

表 12: 被験者のメッセージの修正前と修正後の好感度の客観的項目の比較

修正前	修正後	共感	関心	好意	ポジティブ	親しみやすさ
3 以下	4 以上	45(54.2%)	36(34.6%)	37(56.1%)	40(40.0%)	66(61.7%)
	3 以下	38(45.8%)	68(65.4%)	29(43.9%)	60(60.0%)	41(38.3%)
4 以上	4 以上	33(91.7%)	10(66.7%)	47(88.7%)	17(89.5%)	8(66.7%)
	3 以下	3(8.3%)	5(33.3%)	6(11.3%)	2(10.5%)	4(33.3%)
3 以上	2 以下	3(2.8%)	4(5.9%)	0(0.0%)	3(3.6%)	5(4.9%)

表 13: 「共感」,「好意」,「親しみやすさ」が改善されたメッセージの例

項目	修正前	評価	修正後	評価
共感	ゴルフも気になる	3	ゴルフもいいよね!	4
	スタバ無難かな	3	スタバもいいね! 定番だけど、他に気になる店ある?	4
好意	しつこいよ	2	ちょっとだけ落ち着こうか	4
	うん	3	楽しみだね! どんな夢だったかすぐに教えるよ!	5
親しみ	いいね! どんなニュースが面白い?	3	おっスポーツニュースいいね! どんな話題があるの?	4
	テレビゲームとか麻雀とかかな	3	テレビゲームとかや麻雀を楽しんでるよ! みんなで同じことをするのって最高だよ	5

表 14: 修正案が提示されたが、被験者が好まない話題だったことにより修正がなされなかった例

修正前	修正後
スポーツ観戦は他はもういいかな ところで他に興味あることある?	さっきも言ったけどあまりスポーツに詳しくないから他にはないかな?
試合はちょっと気が進まないかな、	試合やだよー
スポーツ観戦は見えないね。インドア派なので!	スポーツは興味ないね。インドアな性格なので、
うーん、特に大したことはないよ〜。ただ疲れちゃったかな!	え、そんな面白いエピソードはないよ。普通に疲れた。
スポーツにはあんまり興味ないんだよね、他に楽しいことを探そうかな	スポーツに興味は湧かないので、観戦にはいかないかな。
一人で黙々と聞いていたので、野外フェスも行ったことないんだよね	残念ながら野外フェスも。。一人で黙々と聞いていたい

## 5.4 実験結果・考察

「太郎さん (A)」と対話した被験者が送信した修正前のメッセージと、修正後のメッセージの好感度の客観的項目の点数の変化を表 12 に示す。なお、割合の値はすべて修正前の点数を前提とした条件付確率で表されており、例えば「共感」において、修正前の点数が3点以下だったもののうち、修正後に4点以上だったものは58.6%であることを示している。

「共感」,「好意」,「親しみやすさ」については、5-6割近い割合で「3 以下」から「4 以上」に修正されてい

たのに対し、「関心」と「ポジティブ」に関しては3-4割程度の修正にとどまっていた。

「共感」,「好意」,「親しみやすさ」が改善されたメッセージの例を表 13 に示す。「共感」は、ユーザの修正により、「楽しいよね!」や「いいね!」など、相手に共感を表す表現が追加されており、「共感」の点数が向上していることが確認できる。「好意」は、修正前は「しつこいよ」や「うん」など、対話相手に対して好意を示さないようなメッセージだったものが、ユーザの修正により改善されており、「好意」の点数が向上していることが確認できる。「親しみやすさ」は、修正前は「テレビとかで観てる」や「好きです」,「麻雀とかかな」など、表現に固さが見られるメッセージが、ユーザの修正により「最高だよ」,「行くよ」など、柔らかい表現に変わっており、「親しみやすさ」が向上していることが確認できる。

一方で、「関心」,「ポジティブ」については他の項目に比べ修正される割合が少なかった。改善されなかった原因として、被験者が嫌いな話題であったものが多くみられた。これより、被験者が嫌いな話題、興味のない話題になったとき、被験者が好感度の高いメッセージを送信しにくくなり、特に「関心」や「ポジティブ」ではそれが顕著に現れやすいと言える可能性が高い。修正案が提示されたが、被験者が好まない話題だったことにより修正がなされなかった例を表 14 に示す。被験者が嫌いな話題については、修正案で違う内容のメッセージを提示されても、被験者が自らの意見を押し通すことが多く見られた。この結果から、被験者が嫌いな

話題や、チャットボットとの好みの対立によって好感度が改善されなかった可能性が高く、特に「関心」や「ポジティブ」に対して大きく影響していたと考えられる。

## 6 おわりに

ChatGPT を用いた対話相手の好感度改善のシステムについて、メッセージ修正案の提示による対話相手からの好感度改善について検証した。

被験者が送信しようとしているメッセージに対して提案システムを利用した場合と、システムを利用しなかった場合でのチャットボットの好感度の違い、およびユーザーが修正案を利用することによる好感度の向上について、好感度の出力結果と被験者のアンケートの回答から検証した。検証の結果、「共感」、「好意」、「親しみやすさ」については、修正案を提示された被験者による修正で、メッセージの高い割合で改善されていた。一方、一部被験者は、チャットボットとの好みの違いが生じた結果、対話相手の話題に関心を示しにくくなることが確認され、「関心」、「ポジティブ」の改善率の低下に影響したと考えられる。

これらの結果を受けて、ユーザーと対話相手の双方の嗜好が対立する場合の修正案の改善が今後の課題である。これにより、システム利用者と対話相手の嗜好を両方取り入れたメッセージの修正案を提示し、より対話相手の好感度を高めることを目標とする。

## 参考文献

- [1] (URL) <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r06/pdf/n2170000.pdf>, 総務省情報通信白書令和6年版, pp.152–157 (2024)
- [2] Zhao, Z., Fan, W., Li, J., Liu, Y., Mei, X., Wang, Y., ... Li, Q. : Recommender Systems in the Era of Large Language Models (LLMs), IEEE Transactions on Knowledge and Data Engineering, Vol.36, pp.6889 – 6907 (2024)
- [3] Liu, J., Shao, Y., Zhang, P., Li, D., Gu, H., Chen, C., ... Gu, N. : Constructing and Masking Preference Profile with LLMs for Filtering Discomforting Recommendation, arXiv preprint, arXiv:2410.05411, (2024)
- [4] 小林峻也, 萩原将文 : ユーザの嗜好や人間関係を考慮する非タスク指向型対話システム, 人工知能学会論文誌, Vol.31, No.1, pp.DSF-A.1 (2016)
- [5] 杉原穂, 目良和也, 黒澤義明, 竹澤寿幸 : 対話を通じて話者の嗜好情報を学習する音声対話システム, 言語処理学会第23回年次大会発表論文集, Vol.23, pp.226 – 229 (2016)
- [6] 目良和也, 市村匠, 相沢輝昭, 山下利之 : 語の好感度に基づく自然言語発話からの情緒生起手法, 人工知能学会論文誌, Vol.17, No.3, pp.186–195, (2002)
- [7] Horsham, Z., Haydock-Symonds, A., Imada, H., Tai, H. C., Lam, L. W., Lui, S. T., ... Feldman, G. : Does learning more about others impact liking them?: Replication and extension Registered Report of Norton et al.(2007) 's Lure of Ambiguity, OSF (2024)
- [8] 村越琢磨 : 他者に対する理解度の認知が好感度に及ぼす効果, 日本心理学会大会発表論文集, Vol.85, No.7, pp.PI-042-PI-042 (2021)