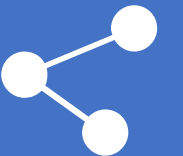


ゼミ 2025.5.20

YANG JIANHUA 202521748



1.論文

1). Large Language Models in Mental Health Care: A Scoping Review (Hua et al.,2024)

i)研究背景と目的

近年、GPT-4などの大規模言語モデル(LLM)の発展により、AIを活用した心理ケアやメンタルヘルス支援が注目されています。本論文は、2019年以降の最新LLMが精神保健分野でどのように利用されているか、その利点・課題・倫理面までを包括的に整理したスコーピングレビューです。

ii)研究方法

PRISMAガイドラインに従い、PubMedやArXivなどの主要データベースから関連文献を系統的に検索し、「T5以降」のLLMを対象に、最終的に34本の研究を分析しています。

iii)技術的な応用と可能性

- ・**共感的な会話エージェント**: LLMは自然で共感的な対話生成が可能となり、感情支援やデジタルカウンセリングに活用されています。
- ・**精神疾患・リスク検出**: うつ病・自殺・不安・認知の歪み等のテキスト自動検出で高い柔軟性と精度を示し、プロンプトのみでも多様な課題に対応できます。
- ・**データ拡張とカスタマイズ**: LLMによる高品質な合成データ生成や、命令型ファインチューニング(Instruction Fine-tuning)を通じた専門用途への適応も可能です。

iv) 主な課題と今後の展望

- ・ **データの代表性と質**：多くのデータがRedditなどSNS由来で臨床的な代表性やラベル基準に課題があります。
- ・ 評価指標の統一性：自動評価と人的評価が混在し、共通フレームワークがなく比較が困難です。
- ・ 倫理・安全リスク：有害な出力、ユーザー依存、プライバシー漏洩等のリスクが未解決です。
- ・ 実臨床応用の障壁：現場での長期有効性や安全性の検証はこれからです。
- ・ **技術的な課題**：
 - a. LLMの“推論”や“共感”の本質的な理解力には限界があり、実際にはパターン記憶や表層的な言語生成にとどまる場合も多い。
 - b. 長文コンテキストの保持、個別化対応、対話の一貫性・多様性の確保などは現在も課題です。
 - c. マルチモーダルやリアルタイム応答性も今後の技術開発テーマです。

v) 結論

LLMは精神保健領域に大きな技術的可能性を持っていますが、「高品質な専門データ」「評価基準の整備」「倫理・法的ガイドライン」が今後の実装には不可欠です。今後は多分野連携を強化し、AI技術を安全かつ実用的に医療現場へ導入することが重要です。

2) Towards Emotional Support Dialog Systems (Liu et al.,2021)

i) データセットの内容と構成

役割と形式データセットのすべての対話は、「求助者 (Help-Seeker)」と「支援者 (Supporter)」による多ラウンドのインタラクションで構成されています。プロのトレーニングを受けたクラウドワーカーが支援者・求助者の役割を演じ、実際の心理支援場面をリアルに再現しています。

ii) 対話内容

各対話は: 平均約30ターン (29.8) の多ラウンド対話で、従来の感情／共感対話データセットよりも多いラウンド数を持ちます。各発話には詳細なラベル付けがあり (支援者は毎ターン必ず具体的な支援戦略を選択)。

対話前:

求助者 (Help-Seeker) の背景ラベル:

・問題タイプ (Problem Type): 学業ストレス、仕事の危機、失恋、友人関係、うつ・気分の落ち込みなど5種類以上。

・経験タイプ (experience_type):

Previous Experience: 同じ悩みを以前にも経験したことがある。

No Experience: 今回が初めての経験。

感情タイプ (Emotion Type): 不安、抑うつ、悲しみ、怒り、恐怖、恥、嫌悪の7種類から選択。

感情の強度 (Emotion Intensity): 1～5段階のスケールで現在の感情の強さを自己評価。

背景説明 (Personal Description): 自由記述形式で、今どんな状況や悩みを抱えているかを具体的に説明。

対話中:

支援者は毎発話ごとに「支援戦略 (Support Strategy)」を7つのうちから選択・記録

求助者は2ターンごとに、支援者の対応に対して5段階で評価 (フィードバック)。

対話後:

求助者は再度、感情強度を記入し、支援者の共感力や話題の関連性を評価。

支援者は求助者の自己開示や説明の詳細さを評価。

iii) 支援戦略(Support Strategies)

各支援者発話は、以下7種類の戦略のいずれか(複数も可)としてラベル付けされています:

Question(質問・問題の探索)

Restatement or Paraphrasing(言い換え・再表現)

Reflection of Feelings(感情の反映)

Self-disclosure(自己開示・似た経験の共有)

Affirmation and Reassurance(肯定・励まし)

Providing Suggestions(提案)

Information(情報提供)

Others(その他)

iv) ラベル付与情報

各ターンの支援戦略。

求助者の感情タイプ・強度(前後比較で「支援効果」を定量化可能)。

問題カテゴリ。

求助者と支援者による詳細なフィードバック。

複数の自動・人的品質管理メカニズムでデータの有効性と信頼性を保証。

```
{
  "dialog": [
    {
      "speaker": "supporter",
      "annotation": {
        "strategy": "Restatement or Paraphrasing"
      },
      "content": "It sounds like you enjoyed the sex a lot, but you also want an emotional connection to a woman."
    },
    {
      "speaker": "seeker",
      "annotation": {
        "feedback": "5"
      },
      "content": "Yes totally. Some of the prostitutes are very sweet, but they don't want anything too long term. Just a weekend at most."
    },
    {
      "speaker": "seeker",
      "annotation": {},
      "content": "Well thank you, you have been very helpful, I feel better now."
    },
    {
      "speaker": "supporter",
      "annotation": {
        "strategy": "Affirmation and Reassurance"
      },
      "content": "Ok, be well."
    }
  ],
  "seeker_question1": "I liked her help.",
  "seeker_question2": "No",
  "supporter_question1": "",
  "supporter_question2": ""
},
{
  "experience_type": "Previous Experience",
  "emotion_type": "anxiety",
  "problem_type": "problems with friends",
  "situation": "not communicative all of a sudden, not wanting to speak over phone too",
  "survey_score": {
    "seeker": {
      "initial_emotion_intensity": "4",
      "empathy": "5",
      "relevance": "5",
      "final_emotion_intensity": "1"
    },
    "supporter": {
      "relevance": "4"
    }
  }
}
```

2. 研究方向

1) 感情支援戦略の予測 (Emotional Support Strategy Prediction)

論文: EmoDynamiX : Emotional Support Dialogue Strategy Prediction
by Modelling MiXed Emotions and Discourse Dynamics (Wan et al., 2025)

「対話戦略予測」を明示的かつ独立したタスクとして扱い、モデルが暗黙的に決定するのではなく、システムの制御性と説明性を向上させることを目指しています。そのために、EmoDynamiXというフレームワークを提案し、ユーザーのきめ細かな感情状態とシステムの支援戦略を動的にモデリングしています。過去の感情分布や戦略選択、発話構造の情報を統合し、「ダミーノード」を導入することで情報の集約とモデルの性能・説明性をさらに高めています。

02 研究方向

2)感情支援応答の生成(Emotional Support Response Generation)

論文:Empathetic Response Generation for Distress Support (Yeh et al.,2023)

本論文は、困難な状況下での共感的対話データセット(RED)を最大規模で構築し、**認知的・情動的共感識別メカニズムを組み合わせた生成モデルを提案しています**。このデータセットを用いて、既存の感情対話モデル(EmoPrependやMEED)および新たに提案した認知的・情動的共感識別メカニズムを備えたモデル(EPIMEED、EPIMEED+)の**4種類の共感応答生成ベースラインモデルを学習させました**。さらに、心理学の観点から「情動的共感」と「認知的共感」を区別し、EPITOMEツールを活用して多層的な共感表現を精緻に識別・定量化し、「コミュニケーション埋め込み (communication embedding)」を生成モデルに導入することで、より豊かな共感的応答を実現しています。

02 研究方向

3)フィードバックに基づく感情支援モデリング (Feedback-based Emotional Support Modeling)

論文:FADO: Feedback-Aware Double COntrolling Network for Emotional Support Conversation (Peng et al.,2023)

FADO (Feedback-Aware Double Controlling Network)というフレームワークを提案し、感情支援対話システムにおける戦略選択と戦略に基づく応答生成を実現しました。本手法では、各ターンごとの「ターンレベルフィードバック」と、会話全体に関する「セッションレベルフィードバック」(ユーザーの感情やストレスの変化)という二重のフィードバック機構を導入し、戦略選択を総合的に最適化することで、ユーザーの現在および全体状態により適合した支援を可能にしています。FADOは、戦略選択の正確性および応答生成の品質の両面で、その時点で最先端(SOTA)の性能を達成しました。

02 研究方向

EmoDynamiXによる戦略予測は、主に対話履歴と感情のダイナミクスに注目していますが、「ユーザー個人差」を十分に考慮していないのが現状です。実際には、同じ問題であっても、ユーザーごとに適切なサポート戦略は異なる場合が多いです。

そこで本研究では、ESConvデータセットに既存するユーザープロフィール情報(経験タイプ、感情タイプ、問題タイプなど)を活用し、各ユーザーによりパーソナライズされた、動的な戦略推薦を実現することを目的とします。

イノベーションのポイント

本研究の最大の特徴は、ESConvの元データに含まれるユーザープロフィール(experience_type、emotion_type、problem_type)を体系的に戦略予測モデルに組み込む点です。これらのプロフィール情報を対話文脈と深く統合することで、モデルがユーザーごとの個人差を自動的に認識し、個々のseekerに最適な動的戦略予測を行うことが可能となります。

今後予定

引き続き関連論文を精読し、必要な知識を習得します。

