

看護過程での臨床判断支援システムにおける 大規模言語モデルを活用したフィードバックの自動生成

佐藤玲央^{*1}, 木口幸子^{*2}, 辻幸美^{*2}, 初見温子^{*2}

藤長すが子^{*2}, 高岡哲子^{*2}, 山川広人^{*3}, 小松川浩^{*3}

*1 公立千歳科学技術大学大学院理工学研究科, *2 北海道文教大学医療保健科学部看護学科

*3 公立千歳科学技術大学理工学部

Automatic Generation of Feedback Utilizing A Large-Scale Language Model in Clinical Decision Support System for Nursing Process

Reo Sato^{*1}, Sachiko Kiguchi^{*2}, Yukimi Tsuji^{*2}, Atsuko Hatsumi^{*2}

Sugako Fujinaga^{*2}, Tetsuko Takaoka^{*2}, Hiroto Yamakawa^{*3}, Hiroshi Komatsugawa^{*3}

*1 Graduate School of Science and Technology, Chitose Institute of Science and Technology

*2 Department of Nursing, Faculty of Health Sciences, Hokkaido Bunkyo University

*3 Faculty of Science and Technology, Chitose Institute of Science and Technology

In basic nursing education, it is required for beginners to gain logical thinking ability applicable in the nursing field. In the previous study, we proposed and developed a clinical decision support system for nursing process, which could not provide feedback information for the next learning considering learners' thinking process and understanding degree. In the present study, we introduce a large-scale language model into generation of the feedback information and examine its effectiveness.

キーワード: 看護過程, 看護基礎教育, 大規模言語モデル, ChatGPT, プロンプトエンジニアリング

1. はじめに

看護師養成機関では、実際の現場が求める問題解決力を基礎とした看護的思考能力を備える人材の育成を、看護基礎教育段階において試みている。看護的思考能力とは、対象患者の全体像を捉え、客観的・主観的データを適切にアセスメント・診断し、看護問題を解決するために立案した看護計画を実施・評価する能力を指す。この能力を用いて看護問題を解決する一連のプロセスを看護過程と呼ぶ[1]が、看護初学者の看護的思考能力を実際の現場が求める水準まで養成することは容易ではない。

看護初学者は、看護的思考能力を座学での事例展開

学習と臨地実習を繰り返すこと通じて身につけていく。事例展開学習に際し、患者事例としてペーパーパシエントを用いるが、学習者の動機付けに重要な患者像や、主体的・能動的な問題解決に重要な臨床場面をいかに現実に近づけ具体化し、看護的思考能力を養成できるかが大きな課題となってきた。

こうした課題に対し、本研究チームでは先行研究において看護的思考能力の向上を狙い、看護実習のプレ教育として看護実習の流れ（シナリオ）に則り、看護的思考過程を自己トレーニングできる臨床判断支援システムを提案してきた。学習者はシナリオに沿って看護過程を思考し、自らの判断で選択をし、それを学習者間で共有することで判断の妥当性を検討する形式と

なっている。この際、指導者からの助言をフィードバック情報として提供されることも重要となるが、先行研究ではそれを考慮したシステムの機能検討までは行っていなかった。

2. 目的

本研究では、先行研究で開発した臨床判断支援システムに対して、学習者の学習履歴とシステム内で設定されている看護シナリオを大規模言語モデルの入力情報として活用することで、指導者のフィードバックの代替になりうるかの検証を行うことを大きな目的とする。そのため、具体的には、対話型の大規模言語モデルである ChatGPT への指示となるプロンプトの設計と生成されるテキストの品質の評価を行うこととした。

3. 臨床判断支援システム

看護的思考過程の確立を狙った臨床判断支援システムは、複数の看護シナリオについて看護過程における「アセスメント」「看護診断」「看護計画立案」の3要素を、それぞれ「課題1：SO 情報分類」「課題2：看護判断」「課題3：看護援助」の3つの課題を図1に示す順序で学習することができる。学習の結果はユーザー毎にIDと紐付けてクラウド上のデータベースに保存する。評価は、選択の確らしさに基づき課題毎に5段階で提示する。本研究では、課題1, 2, 3に対する学習者が回答した理解度と選択する根拠などの情報を入力プロンプトに採用した(以下成績情報という)。また、システム上で設定された看護シナリオ(患者の情報)と各課題に対してシステム上で設定されている指導側のポリシーについてもプロンプトに入力することとした。これにより、本システム上での学びの結果か

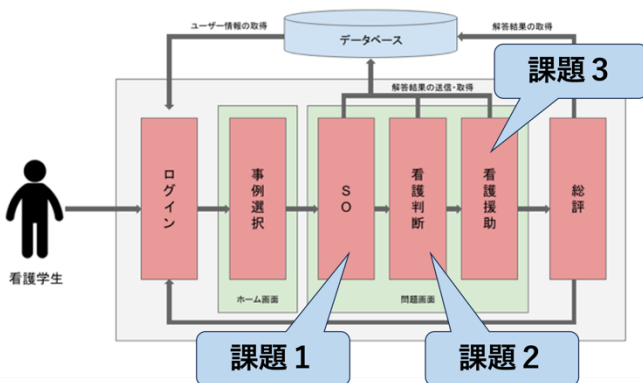


図1 システム構成図

ら得られるデータのみで、どこまで自動フィードバックを返すことができるかを検証することとした。なお、指導メッセージを出すために、ChatGPTに自らが指導役であることを示し、また指導する観点についても明示し、これに沿ったメッセージを返すようにした。

4. ChatGPT の導入

4.1 ChatGPT

OpenAI社が2022年に公開したChatGPT[2]は、対話型の大規模言語処理モデルである。ユーザーとのインタラクションを通じて、質問に回答したり情報を提供したりすることができる。ChatGPTは高度な言語理解能力と文章生成能力を備えており、応用的な活用が期待されている[3]。応用的な活用事例として、阿部らの研究[4]では、薬についての相談ができるChatGPTを組み込んだアプリケーションを開発し、セルフメディケーション実践へのChatGPTの介入について考察している。

4.2 プロンプト

ChatGPTへの指示としてプロンプトを入力する。プロンプトは大規模言語処理モデルに対して任意のタスクを実行させるための指示であり、適切なテキストの生成には、ユースケースに応じた適切なプロンプトの設計が求められる。

4.3 本研究におけるChatGPTの使用方法

本研究では、臨床判断支援システムでのフィードバックの自動生成を見据え、システムのデータベースに蓄積される成績データと看護シナリオを基にプロンプトを設計し、ChatGPTによるフィードバックの生成を試みる。プロンプトの設計については、DAIR.AIがオープンソースとして公開しているPrompt Engineering Guide[5]を参照し、学習者の思考過程と理解度を判別するプロンプトを設計する。

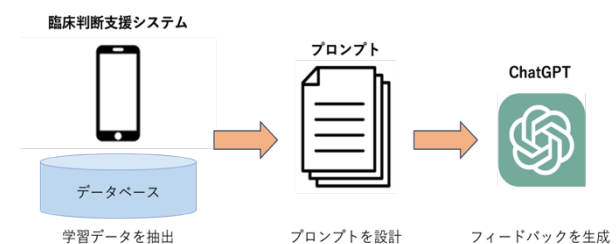


図2 ChatGPT の使用順序

4.4 プロンプトの構成

プロンプトは、ChatGPT に課す「役割」、達成したい内容の「指示」、生成するフィードバックの「フォーマット」、看護シナリオに関する情報・各課題の評価基準となる情報・学習者の成績データの3つの「文脈情報」を基本的な構成とする。実際に設計したプロンプトの一部を表5に示す。

5. 評価

5.1 評価対象

「胃がん」「肺炎」「脳梗塞」「糖尿病」の4症例の看護シナリオに対して、それぞれ成績のダミーデータを3人分作成し、合計12人分のプロンプトを設計した上、ChatGPT に生成させたフィードバックの文章を評価対象とする。実施に生成されたフィードバックの文章の一部を表6に示す。

5.2 評価指標

表1に示す項目について、生成したフィードバックの文章を臨床判断支援システムの課題1から課題3それぞれに対して評価する。成績データが反映されているかについての項目は「A:反映されている, B:反映されていない」で評価する。5段階評価が適切に判別しているかについての項目は「A:判別している, B:判別していない」で評価する。具体的な改善アドバイスを含んでいるかについての項目は「A:含んでいる, B:含んでいない」で評価する。

表1 評価項目

	項目	評価
1	成績データが反映されている	A.反映されている B.反映されていない
2	5段階評価を適切に判別している	A.判別している B.判別していない
3	具体的な改善アドバイスを含んでいる	A.含んでいる B.含んでいない

5.3 結果

5.2節の評価指標に基づく評価結果を項目毎に示す。表2は成績データの反映状況の項目、表3は5段階評価の判別状況の項目、表4は具体的なアドバイスの項目を示している。

成績データの反映状況については、課題1で58%、

課題2と課題3で92%反映されており、成績データを反映していない部分については、「あなたの成績はまだ評価されていませんが、取り組みに真剣さを感じられます。」などのフィードバックが散見された。

5段階評価の判別状況については、課題1で42%、課題2で75%、課題3で92%の精度で判別しており、適切に5段階評価が判別できていない部分については、最高評価を獲得しているにも関わらず、「課題2の患者問題の判断に関する課題については、一部改善の余地があります。」などのフィードバックが生成された。

具体的なアドバイスの項目については、課題1で0%、課題2と課題3で33%の精度で具体例を含む改善アドバイスが生成された。生成された具体的な改善アドバイスには「より良い援助内容を選択するためには、患者の状態やニーズに基づいて適切な援助計画を立案することが大切です。患者の意向や疲労度を考慮し、より良い援助内容を提案できるように工夫しましょう。」など、思考過程の手がかりを示していた。

総じてChatGPTは、精度の良し悪しとは別に、12人に対して異なる表現でかつ観点毎に前向きなフィードバックを生成しており、指導メッセージとして自然な文章となっている。

表2 成績データの反映状況

	課題1	課題2	課題3
Aの件数	7	11	11
Bの件数	5	1	1
Aの割合	58%	92%	92%
Bの割合	42%	8%	8%

表3 5段階評価の判別状況

	課題1	課題2	課題3
Aの件数	5	9	11
Bの件数	7	3	1
Aの割合	42%	75%	92%
Bの割合	58%	25%	8%

表4 アドバイスの具体性

	課題1	課題2	課題3
Aの件数	0	4	4
Bの件数	12	8	8
Aの割合	0%	33%	33%
Bの割合	100%	67%	67%

表 5 設計したプロンプトの一例

【役割】
<p>### 役割 ### あなた(GPT)の役割を「看護学初学者へ看護過程に沿ったアドバイスをする教員」とします。</p> <p>アドバイスの方向性は「学習者が「患者情報を適切にアセスメントする」「患者問題をエビデンスに基づいて判断する」「患者問題に対して適切な援助計画を立案する」の3つを看護過程に沿って実施できるようにするために、エビデンスに基づいた援助計画の立案に関するアドバイスをする」とします。</p>
【指示】
<p>### 指示 ### アドバイスをするための事前情報として「患者情報」と「患者問題を判断するエビデンスの重要度の情報」と「患者問題に対する援助内容の重要度の情報」と「学習者の成績の情報」があります。</p> <p>「患者情報」は、症例、氏名、年齢、性別、経過の情報が含まれています。</p> <p>「患者問題を判断するエビデンスの重要度の情報」は、エビデンスの一覧であり、重要度を4段階（3:重要である, 2:どちらかというとも重要である, 1:どちらかというとも重要ではない, 0:重要ではない）で示しています。</p> <p>「患者問題に対する援助内容の重要度の情報」は、援助内容の一覧であり、重要度を4段階（3:重要である, 2:どちらかというとも重要である, 1:どちらかというとも重要ではない, 0:重要ではない）で示しています。</p> <p>「学習者の成績の情報」は課題1、課題2、課題3それぞれの評価と、各課題の評価を平均した総評の情報が含まれています。</p> <p>学習者は課題1「アセスメントに関する課題」、課題2「患者問題の判断に関する課題」、課題3「援助内容に関する課題」に取り組みます。</p> <p>エビデンスに基づいた患者問題の判断と適切な援助内容を適切に選択できているかを評価し、「看護過程に沿った適切な看護計画をエビデンスに基づいてどのように立案すれば良いか？」を学習者にアドバイスしてください。</p> <p>アドバイスする際には、学習者を傷つけないようにポジティブなメッセージでアドバイスしてください。</p>
【フォーマット】
<p>### アドバイス文のフォーマット ### 【はじめに】<文章量は2文程度で、良いところを褒めるメッセージを書いてください。>【課題1に対する評価】<文章量は1文程度で、学習者の成績の情報から学習者に現状を伝えてください。></p> <p>【課題2に対する評価】<文章量は2文程度で、学習者の成績の情報から学習者のエビデンスの選択を評価し、より良いエビデンスの選択を行うためのアドバイスをしてください。></p> <p>【課題3に対する評価】<文章量は2文程度で、学習者の成績の情報から学習者の援助内容の選択を評価し、より良い援助内容の選択を行うためのアドバイスをしてください。></p> <p>【全体を通しての評価】<文章量は2文程度で、学習者の成績の情報から学習者の現状と今後の伸び代を伝えてください。></p>
【文脈情報（看護シナリオに関する情報）】
<p>### 患者情報 ### =====</p> <ul style="list-style-type: none"> ・症例：胃がん ・氏名：武川浩 ・年齢：55歳 ・性別：男 ・経過：年1回、職場の健康診断の他、胃がんの検診のため、胃カメラの検査を受けていた。胃カメラにてポリープが見つかり、精査を進められた。近医を受診し胃カメラにて組織片を採取し、生検の結果、早期胃がんと診断され、内視鏡的切除術による治療目的で入院となった。 <p>=====</p>
【文脈情報（各課題の評価基準となる情報）】
<p>### 患者問題を判断するエビデンスの重要度の情報 ### =====</p> <ul style="list-style-type: none"> ・喫煙歴20年、1日10本喫煙（重要度：3） ・「長年タバコを吸っている」（重要度：3） ・外来受診時肺機能検査 %肺活量80以上、1秒率65（重要度：3） ・医師の所見 肺機能検査にて1秒率において若干の低下があり（重要度：3） ~~~~~（中略）~~~~~ ・「来週から仕事に戻るかな」（重要度：0） ・術後7日目：退院（重要度：0） ・検査の結果：転移の可能性はなく、抗がん剤の使用もしなくても良いでしょう（重要度：0） ・「先生の話聞いて、これですっきりした気持ちで仕事に戻れます。」（重要度：0） <p>=====</p> <p>### 患者問題に対する援助内容の重要度の情報 ###=====</p> <ul style="list-style-type: none"> ・肺合併症のリスク（重要度：3） ・廃用症候群のリスク（重要度：3） ・肺塞栓症のリスク（重要度：3） ・疼痛（重要度：2） ~~~~~（中略）~~~~~ ・胃痛の軽減（重要度：0） ・腸蠕動運動の観察（重要度：0） ・臥床状態による肺炎のリスク（重要度：0） ・痰貯留のリスク（重要度：0） <p>=====</p>

【文脈情報（学習者の成績データ）】
<p>### 学習者の名前 ### 山下康太</p> <p>### 学習者の成績の情報 ###</p> <p>=====</p> <ul style="list-style-type: none"> ・課題1. アセスメントに関する課題の評価：5（1~5の段階評価。数値が大きいほど良い。） ・課題2. 患者問題の判断に関する課題の評価：5（1~5の段階評価。数値が大きいほど良い。） ・課題2において選択した患者問題：内視鏡下胃がん摘出術後の合併症のリスク（呼吸器・循環器） ・課題2において患者問題のエビデンスとして選択した項目の一覧：喫煙歴20年、1日10本喫煙、「長年タバコを吸っている」、外来受診時肺機能検査 %肺活量80以上、1秒率65、医師の所見 肺機能検査にて1秒率において若干の低下があり、喫煙にて術後、痰が多くなるなどの悪影響、「煙草は悪いことはわかっているがストレスで止められない」、病室入室後体温36.8℃、病室入室後脈拍80回/分、病室入室後血圧128/78mmHg、病室入室後呼吸17回/分 ・課題2においてエビデンスを選択した理由：患者の喫煙歴が長く、呼吸器疾患を引き起こすリスクが通常に比べて高いと考えたため。 ・課題3. 援助内容に関する課題の評価：5（1~5の段階評価。数値が大きいほど良い。） ・課題3において援助内容として選択した項目の一覧：肺合併症のリスク、廃用症候群のリスク、肺塞栓症のリスク ・課題3において援助内容を選択した理由：術後数日は合併症の発症リスクが高く、経過観察を通してリスク管理を行うことが重要と考えるため。 ・総評：5（課題1, 課題2, 課題3の平均値） <p>=====</p>

表 6 フィードバックの一例

【生成されたフィードバック】
<p>【はじめに】 山下康太さん、あなたの成績についてお話しいたします。まず、取り組んだ課題については素晴らしい取り組みを見せていただきました。</p> <p>【課題1に対する評価】 課題1のアセスメントに関する課題については、評価できる成果をお見受けしました。患者の情報に基づいて適切なアセスメントを行うことができましたね。</p> <p>【課題2に対する評価】 課題2の患者問題の判断に関する課題については、一部改善の余地があります。エビデンスの選択において、より重要な情報にフォーカスすることが求められます。具体的なアドバイスとしては、喫煙歴や肺機能検査結果など、喫煙に関連するエビデンスを重視すると良いでしょう。</p> <p>【課題3に対する評価】 課題3の援助内容に関する課題については、十分な成果を上げていただきました。適切な援助内容を選択する能力を持っていることが分かりました。</p> <p>【全体を通しての評価】 山下康太さんの成績は非常に優秀です。課題1では適切なアセスメントを行い、課題3では適切な援助内容を選択することができました。ただし、課題2においてはより重要なエビデンスの選択に注意が必要です。今後も学習を継続し、エビデンスに基づいた看護計画の立案能力を高めていくことが大切です。頑張ってください！</p>

6. まとめ

参考文献

本研究では、ChatGPT への指示となるプロンプトの設計と生成されるテキストの品質の評価を行った。

評価結果から、成績情報を反映したフィードバックが可能であることが明らかとなった。しかし、表 2 から表 4 に示す課題 1 の評価の精度は、十分な精度が確認できなかった。要因として、プロンプトの総量に対して、課題 1 の成績データに関連する情報の割合が非常に少ないため、ChatGPT が適切に文脈を捉えることができなかったと考えられる。また、表 4 に示す具体的な改善アドバイスについても十分な精度が全ての課題を通じて確認できなかった。要因として、フォーマットで指定したテキスト分量の範囲では、具体的に改善方法を示すことが出来なかったと考えられる。

今後、精度の向上に向け、プロンプトを見直し追加の検証を実施する予定である。

- (1) 東サトエ：「看護基礎教育における看護過程の理論モデルと教育プログラムの検討」, 南九州看護研究誌 (2017)
- (2) OpenAI: "Introducing ChatGPT"
<https://openai.com/blog/chatgpt>
- (3) 武田俊之: "生成系 AI の教育および教育研究へのインパクトについて", 教育システム情報学会研究報告, Vol.38, No.1, pp.51-56 (2023)
- (4) 阿部祐来, 南野早芸, 真嶋由貴恵, 榎田聖子, 川原淳: "大学生のセルフメディケーション実践支援教育に向けた LINE 版おくすり手帳の開発", 教育システム情報学会研究報告, Vol.38, No.1, pp.17-21 (2023)
- (5) DAIR.AI: "Prompt Engineering Guide"
<https://www.promptinguide.ai/jp>