

音声による構造化介護記録アプリのための データ収集手法の提案

田中 龍之介¹ Tittaya Mairittha¹ 井上 創造¹

概要：本稿では、スマートフォンをベースとした Dialogue System Care Record を拡張して、介護記録のアプリのためのコーパス収集・拡充の方法と実験提案を行なった。日本語コーパスとして NTCIR-13MedWeb を使用しているため、拡充するデータとしてテキストと8つの病気ラベルを保存する。拡充したデータを使用してモデルの再構築を行うことで、モデル制度の向上が図ると想定され、現場の介護士が介護記録を取りながら、コーパス作成が行える。

A Data Collection Method for Structured Care Record Application with Audio Input

RYUNOSUKE TANAKA¹ Tittaya Mairittha¹ SOZO INOUE¹

1. 初めに

近年の高齢化社会の加速の中、高齢者を対象とした介護サービスの需要が増加している。需要が増加している一方で、人材不足や介護業務の多さなどにより、介護の現場は逼迫している状態が続いている。介護業界の問題として介護記録の難しさが挙げられる。介護記録をつけることは必至の業務である一方で、介護記録のフォーマットや書き方の指導については現場に一任されている。よって、介護士が介護業務を書くための指導などの介護以外の業務により業務自体が圧迫されているのが現状である。

NTT データ研究所の調査 [1] では、介護の現場において介護記録のフォーマット統一の必要性や問題点について言及されている。解決する手法として、介護記録の電子化などの手法が提案され、様々な研究が行われている。同研究室においても、介護記録の電子化や行動認識を複合的に使用した、能動的介護記録作成手法や英語での対話的介護記録作成の手法が提案されている。本項では、日本語での対話的介護記録作成を目的としたシステム構築のための問題点である、日本語でのコーパスが少ないといった問題に注目して、効率的で介護士に負担をかけない形で日本語での介護コーパスを作成する手法及び、実験計画を提案する。

システムの活用として、介護記録とコーパス作成をスムーズに行うことが可能になる。

2. 先行研究・関連研究

先行研究として、対話的介護記録システムに関する研究と日本語での対話的介護記録システム構築における問題について紹介する。

2.1 対話的介護記録システム

対話的介護記録システムは、同研究室 Tittaya ら [2] が提案している Dialogue System Care Record (以下 DSCR) を参考にしている。Tittaya らが提案しているシステムでは、スマートフォンをベースとして音声入力を使用して効率的に介護記録を入力するシステムである。音声認識を利用したシステムでは、従来の電子カルテよりも文章加速度が向上し、満足度に関しても上昇することが報告されている。DSCR は、Natural Language Understanding (以下 NLU) のタスク型思考対話システムである。そのため、DSCR 構築のためにはタスクに見合ったコーパスが必要となる。Tittaya らが提案しているシステムでは英語のコーパスを使用しているため、英語の DSCR となっている。

¹ 九州工業大学大学院

2.2 医療・介護における NLU

臨床記録上で実行されるテキスト分類や NLU は、不均衡なデータセット、略語の曖昧な性質、意味上の曖昧さなど、パブリックドメインと比較して特有の課題がある [3]. 加えて、日本では個人情報の観点から臨床記録といったデータが公開されることが希少である。その一方で、臨床記録を使用して臨床的な意思決定やコホートの特定を行う研究が行われている。例として、Uzuner らは [4], Partners HealthCare を元に患者の喫煙状況を特定しました。また、Uzuner らは、患者の退院時サマリーから、肥満とその最も一般的な 15 の併存疾患に関する情報を分類・抽出を行っている。Meystre らは [5] は、ナラティブな臨床試験の説明書から抽出した情報を対応する電子カルテにリンクさせ、患者をケアする臨床医に警告することで、乳がん臨床試験のサンプルに対する患者の適格性を特定している。このように、医療・介護の現場においてテキストなどの情報から患者や医師に支援を行う研究が多くなされている。上記の研究の特徴として英語の臨床記録や電子カルテの情報をを用いて患者の治療ないし、医師の治療に通知を行う側面が強いことが挙げられる。

2.3 日本語 DSCR の問題点

日本語での DSCR [6] (以降 既存のシステム) は前節の DSCR をベースにコーパスとして若宮ら [7] が公開している NTCIR-13 MedWeb を利用して構築されている。NTCIR-13 MedWeb では日英中のテキストに 8 つのマルチラベルが付与されている。考察によって誤差には 2 つの要因を挙げている。

- 音声入力時の誤変換による誤判定
- NLU モデルによる誤判定

NLU モデル作成時と実験時の精度が大きくずれた要因として、NLU モデルによる誤判定が大きな要因だと挙げられている。NLU モデルによる誤判定の要因としては、コーパス内に含まれない表現を用いた発話が挙げられており、コーパスを拡充することで改善できると考えられている。

コーパス作成の手順としては、アウトソーシングによる作成やインターネットのスクレイピングなどを用いる方法などが検討されている。アウトソーシングによる作成は膨大な費用と時間を要する。インターネットのスクレイピングを用いる方法はコーパスのラベリング制度が NLU の制度に直結するため収集したデータを再確認し、修正する必要があると考えられる。コーパスという特徴上、システムの品質にはデータの品質が直結するため、使用上のシナリオが介護記録に用いるコーパスであれば、介護士に作成を依頼する手法が最も効果的であると考えられる。

3. 要件定義

前章での問題点を考慮して、介護記録を作成しながらコーパス収集を行えるように改良する日本語 DSCR の要件定義を行う。作成するシステムでは、Fonlog [2][6] の DSCR を改良する形でシステム構築を行う。

3.1 音声による構造化介護記録アプリの要件 介護士の負担にならないシステム

大前提として、介護士の負担を減らすことが目的である。そのため、システム自体が、介護の現場に負荷をかけないような設計が必要になる。理想とするシステムは、介護記録システムと連動したコーパス収集システムが理想である。システムを使用する対象は、現場の介護士を対象とする。

導入の容易さ

介護の逼迫した現場を考慮して、介護士自体のコストだけでなく介護施設に対してのコストを考える必要がある。システム自体のコストが低いほど導入が容易になると考える。既存のシステムでは、スマートフォンをベースとしたシステムになっており、インターネット環境と記録する人数分のスマートフォン、NLU サーバ・DSCR サーバが必要になる。動作性の観点からは、単純な操作とシンプルな画面設計が必要となる。コスト的観点からでは、インターネットが導入していない施設に関しては、オフラインで全てが完結するシステムが好ましいと考える。導入時のシステム要点に関しては、導入する対象ごとに再度検討する必要があると考える。

3.2 データ収集機能のための要件 コーパス収集のための再評価システム

システムの目的としては、現場の介護士にコーパス作成の助力を願うことである。既存のシステム構成では、判定した結果を表示することは可能である。だが、再評価することができない。既存のシステムに再評価する仕組みを追加し、サーバに保存することで現場の介護士が使用しながら、アノテーション済みのデータを収集する事ができる。

収集時の誤差要因判別

前章で触れた問題点では、誤差の要因には音声入力時の誤変換と NLU モデルによる誤判定の 2 つの要因がある。音声入力時の誤変換による誤判定があるため、システム自体では発話された音声为正しく認識されているかどうかを検知する必要がある。また、モデルが予測した結果が正しいかどうかの判定についてもシステムでは判別することができないため、正しい判定かどうかについても検知する必要がある。コーパスという特徴上、収集するデータはある一定の品質を担保する必要がある。そのため、どちらの要因もしくは両方の要因によって再評価を実行したかを判別

する必要がある。

3.3 目的別要件定義

理想的なシステム要件定義

理想とするシステム自体の要件定義をまとめる。条件として、インターネット設備などは導入済みであり、ランニングコストをある程度確保できているものとする。

- 介護記録
- 容易な操作性
- 再評価システム
- 最適なコスト設計
- 発話からの文字起こし正誤判定
- モデル予測の正誤判定

本稿でのシステム要件定義

本項での目的は、DSCR とコーパス収集の両方が備わっているシステム手法の提案である。よって、既存のシステムと同等のシステムを用い、既存のシステムで用いたコーパスである NTCIR-13 MedWeb を拡充するシステム設計を行う。前提条件として、インターネット設備などは導入済みであり、サーバなどのランニングコストは十分にあるものとする。

- 容易な操作性
- 再評価システム
- 発話からの文字起こし正誤判定
- モデル予測の正誤判定

4. 実装内容

本稿で作成したシステムは、実際の介護現場に導入前の専攻実験として、コーパスの拡充が可能かどうかについて検討材料として実装した。

4.1 システム構築

実装内容について図 1 を用いて説明する。本稿のシステ

ムは、DSCR である Fonlog, Fonlog サーバ, NLU サーバの 3 つで構成されている。図 1 の 1 では、使用者が Fonlog に入力を行う。本稿のシステムは DSCR であるため、初期の入力は発話を用いる。その後、図 1 の 2 で Fonlog から発話からの文字起こしテキストが NLU サーバに送信される。NLU サーバからは、モデルが予測したラベルが返答される。使用者は Fonlog 上で表示された結果を確認後、再評価もしくはフラグの入力を行う。図 1 の 3 では Fonlog が Fonlog サーバとの通信を通して、個人の識別及びデータの保存を行なっている。

Fonlog

Fonlog は DSCR であり、スマートフォンベースのシステムである。基本的な機能としては、Fonlog サーバと連携することで使用者ごとの記録を行うことができる。個人の情報や入力情報に関しては、スマートフォンのローカルストレージに一時的に保存されているが、全て API を通してサーバ上に保存される。発話されたテキストに関しては、NLU サーバに送信され、モデル予測を取得したのちに、ローカルストレージに一時的に保存する。

NLU サーバ

NLU サーバは既存のシステムと同等で TensorFlow[8] で NLU モデルを作成しており、API サーバとして Flask[9]

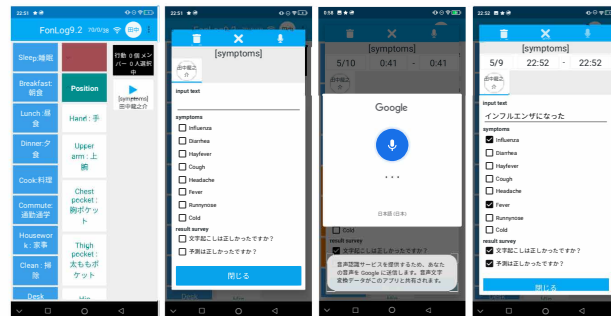


図 2 入力シナリオ

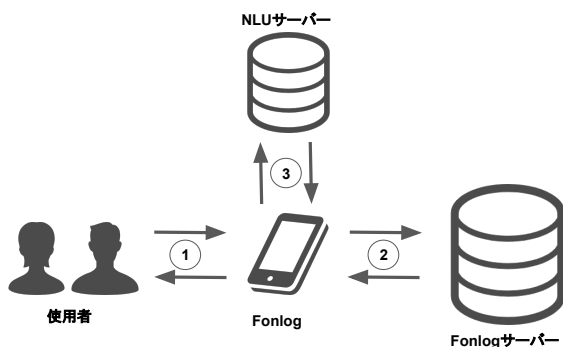


図 1 システム構成図

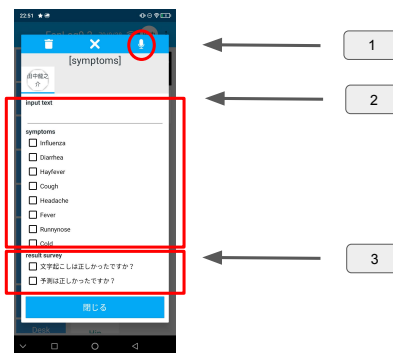


図 3 画面説明図

を使用している。コーパス収集のためにサーバ上で Fonlog から送信されたユーザ ID とテキスト、モデル予測結果が保存される。モデル予測結果ではコーパスである NTCIR-13 MedWeb と同様のラベルである 8 つのラベルが True -False 形式で保存される。

Fonlog サーバ

Fonlog サーバには、Fonlog 上で入力したすべての情報と個人識別のためのユーザー情報が保存される。NLU サーバから Fonlog 上に情報が送信され、情報の再入力を行なったデータが Fonlog サーバで保存される。保存される情報としては、使用者の識別情報、編集されたテキスト及びラベル、発話からの文字起こし正誤判定、モデル予測の正誤判定などが保存される。

4.2 入力シナリオ

入力シナリオについて、図 2 を用いて説明する。主要な画面は 2 つである。まず、一番左端が初期画面である。初期画面では、3 つのラインで画面が構成されており、記録したい情報を選択することで記録ボックスを作成する。記録ボックスを作成したのちに、詳細入力を行う場合では、長押しすることで図 2 の左から 2 番目の詳細ボックスが表示される。詳細ボックスの説明は次節で行う。詳細ボックスでは、DSCR を使用することができ入力が終了すると詳細ボックスを閉じることで情報入力が終了する。

4.3 詳細入力

DSCR の詳細入力画面について図 3 を用いて説明する。詳細入力画面では上部のヘッダー、入力画面、閉じるボタンの 3 つで構成されている。図 3 の 1 を入力することで、DSCR シーケンスに移行する。発話終了がトリガーとなり、NLU サーバに情報が送信される。NLU サーバからの返答の結果は、図 3 の 2 の欄に自動で入力される。もし情報の再入力が必要だった場合は、図 3 の 2 を直接編集する。図 3 の 3 では文字起こし正誤判定、モデル予測の正誤判定がそれぞれフラグとして入力することができる。

4.4 要件の確認

作成したシステムが要件を満たしているかどうかを照らし合わせる。

容易な操作性

主に使用する画面を 2 画面と極力少なくし、DSCR をマイクのアイコンに変えることで視覚的にわかりやすい画面設計を行なった。また、True -False の入力部分では Checkbox を採用することで、容易な入力を意識した。

再評価システム

NLU サーバからの返信を表示させ、テキスト及びラベルの判定をそれぞれ変更できるように画面とシステムを設計した。

発話からの文字起こし正誤判定

発話からの文字起こしが正しい場合は True を正しくない場合は False を入れるような Checkbox を作成した。

モデル予測の正誤判定

NLU モデルの予測が正しい場合は True を正しくない場合は False を入れるような Checkbox を作成した。

5. 実験計画

5.1 実験定義

本項での目的は、効率的に日本語の介護コーパスを収集する手法の提案である。作成したシステムは、3 章で定めた以下の要件を満たしている必要がある。

- 容易な操作性
- 再評価システム
- 発話からの文字起こし正誤判定
- モデル予測の正誤判定

上記の条件を満たしていることを確認するための実験をそれぞれ定義する。

容易な操作性

DSCR を構築することで、使用者の満足度や入力速度が向上することが報告されているが、システム全体を通して満足度や入力速度が低下することは避ける必要があるため、システムの改良を行なった場合では確認することが好ましい。確認する方法としては、手入力での情報入力と DSCR を用いた情報入力をした場合での時間計測とアンケートによる評価を用いる。

再評価システム

本稿の目的の一つとしては、コーパス収集手法の提案である。再評価システムを用いてデータ収集を行い、コーパスの拡充を行う。コーパスの拡充によって、特定のコミュニティ内に強いコーパスが作成することができると考える。結果的に NLU モデルの実証実験での精度が向上すると考える。確認する手法として、既存システムの NLU モデルと拡充したコーパスの NLU モデルの精度比較を行う。

発話からの文字起こし正誤判定

発話からの文字起こしの正誤判定を確認することで誤差の要因割合を算出することができると考える。また、満足度などのアンケートと相互比較をすることで使用者の満足度と関係のある誤差要因が割り出せると考える。

モデル予測の正誤判定

モデル予測の正誤判定を確認することで誤差の要因割合を算出することができると考える。発話からの文字起こし正誤判定と同様で、満足度などのアンケートと相互比較をすることで使用者の満足度と関係のある誤差要因が割り出せると考える。

5.2 実験内容

前節の実験定義を踏まえて、2つの実験に分割する。

コーパス拡充のためのデータ収集実験

以下の手順に従って実験を行う。

- 8つのラベルを発話者に開示する。
- 発話者にラベルを決定してもらい、DSCR に対して入力を行う。
- DSCR の出力に対して再評価を入力する。

DSCR アンケート

以下の手順に従って実験を行う。実験はテキストごとに録画を行い、入力開始から入力終了までの時間を測定する。入力では手入力と DSCR を用いた入力のそれぞれのパターンで入力してもらう。

- 使用者にテキストを開示する。
- システムに入力を行なってもらう。

実験終了後、入力に関しての満足度アンケートを実施する。アンケート内容は以下の5つを五段階評価で行う。

- システムはわかりやすかったですか？
- システムはあなたの言ったことを理解しましたか？
- 必要な情報を記録するのは簡単でしたか？
- システムとのやり取りの速度は適切でしたか？
- システムの入力は簡単でしたか？

6. 終わりに

本稿では、DSCR に加えてコーパスを収集する手法と実験計画について提案した。本稿のシステムでは、NLU モデルには目的に則したコーパスが必要である点を考慮して NTCIR-13 MedWeb を拡充する仕組みである。NTCIR-13 MedWeb を拡充することが可能になれば、他のシチュエーションにおいても、初期に少ないデータセットを作成し、利用しながらデータセットを拡充することが可能になる。

参考文献

- [1] 株式会社NTTデータ経営研究所 ”介護記録法の標準化に向けた 調査研究事業 報告書”, 2020/3.
- [2] Tittaya Mairittha , Nattaya Mairittha, 井上 創造: ”Evaluating a Spoken Dialogue System for Recording Systems of Nursing Care”, MDPI Sensors, Vol. 19, No. 3736, pp. 12 pages, 2019/08/29.
- [3] Ghulam Mujtaba, Liyana Shuib, Norisma Idris, Wai Lam Hoo, Ram Gopal Raj, Kamran Khawaja, Khairunisa Shaikh, Henry Friday Nweke, ”Clinical text classification research trends: Systematic literature review and open issues, Expert Systems with Applications”, Volume 116, 2019, Pages 494-520,
- [4] Özlem Uzuner, PhD, Ira Goldstein, MBA, Yuan Luo, MS, Isaac Kohane, MD, PhD, ”Identifying Patient Smoking Status from Medical Discharge Records”, Journal of the American Medical Informatics Association, Volume 15, Issue 1, January 2008, Pages 14–24, <https://doi.org/10.1197/jamia.M2408>
- [5] Meystre SM, Heider PM, Kim Y, Aruch DB, Britten CD. ”Automatic trial eligibility surveillance based on unstructured clinical data.” Int J Med Inform. 2019 Sep;129:13-19. doi: 10.1016/j.ijmedinf.2019.05.018. Epub 2019 May 23. PMID: 31445247; PMCID: PMC6717538.
- [6] 田中龍之介 Tittaya Mairittha , 井上 創造: ”自然言語を用いた構造化介護記録アプリの評価”, SOFT 九州支部 学術講演会, pp. 5 pages, 2020/11/28.
- [7] Shoko Wakamiya, Mizuki Morita, Yoshinobu Kano, Tomoko Ohkuma and Eiji Aramaki: Overview of the NTCIR-13 MedWeb Task, In Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR-13), pp. 40-49, 2017.
- [8] TensorFlow <https://www.tensorflow.org/overview?hl=ja>
- [9] Flask <https://flask.palletsprojects.com/en/1.1.x/>