

論文内容の要旨

博士論文題目

A Study on Clinical Natural Language Processing Covering Electronic
Medical Records and Patient Narratives

氏名 柴田大作

(論文内容の要旨)

(1, 200字程度)

自然言語処理で、電子的診療録と患者発話データを利活用する研究である。医療分野における自然言語処理において、研究に使用されるデータは大別して2種類ある。1つ目は、電子的診療録 (Electronic Medical Record: EMR)である。EMRは、検査値やコードなど構造的な情報も多いが、臨床的に重要な記述について

は、診療録、看護記録や読影レポートにおける非構造的なテキストとして表現されており、利活用が困難なデータであった。もう1つが、発話データ、特に患者と医療者との対話や患者とその家族のコミュニケーションなどの患者発話を含むデータである。医師不足や超高齢化社会など医療リソースに関する問題が多い日本において、患者データを医療に活かす技術の開発は喫緊の課題である。このような2つの問題について、これまで、これらの両方を扱い体系だった研究を行った事例はなく、本研究は、はじめて一貫した枠組みで両者を扱ったものである。具体的には、以下の3つの研究から構成されている。

まず、EMRに出現する疼痛表現の事実性の分析である。疼痛に関する情報はEMRにフリーテキストとして記載されるため、さまざまな表現で記載されるが、これを区別して抽出する手法を提案した。研究では、疼痛表現のあり方について、15種類の事実性（疼痛があったか、なかったか、疑われていたか）が存在することを確認し、さらに、文脈を考慮することができる埋め込み表現を用いることにより、高い精度で疼痛表現を抽出することに成功した。

2つ目の研究は、読影レポートを構造化された表形式に変換するシステムの開発である。読影レポートは放射線科医によって記載されるが、自然文であり、

臨床研究や医療人工知能の開発などに利活用することが困難であった。そこで、患者情報を表形式に変換することで、再利用が容易な形にした。

3つ目の研究は、アルツハイマー型認知症のスクリーニングのための、発話書き起こしからの認知機能測定手法の確立である。近年、認知症患者の数が世界中で増加しており、その中でもアルツハイマー型認知症が全体の67%を占めると言われている。アルツハイマー型認知症の根本的な治療方法は確立されていないが、薬物治療などで症状の進行を遅らせることは可能であるため、アルツハイマー型認知症の早期発見は重要な課題となっている。これを極めて容易かつ安価に行うため、言語による測定を試みた。近年、いくつかの研究グループが類似した研究を行っているが、基準となるデータがなく、比較が困難であった。本研究はデータを独自に収録、構築することから開始し、さらに認知機能測定の研究までを行った。

氏名	柴田大作
----	------

(論文審査結果の要旨)

本論文は、医療言語処理に関する包括的な論文で、患者データの解析、カルテ（読影レポート、看護記録）の解析、などさまざまな医療テキスト解析をまとめた研究である。

本邦において、該当分野（医療言語処理研究）に研究は少なく、意欲的な研究である。特に以下の点について貢献が大きい。

まず、介護記録というまだ、情報処理が十分適応されていない分野にて、はじめて本格的な自然言語処理を導入した点。特に、もっとも頻用される表現である<疼痛>に注目し、その表現を吸収しているが、他の表現にも適応可能な有用性の高い研究である。

読影レポートの変換システム開発については、解析システムは多く提案されているが、いずれも要素システムであり、はじめてこれをまとめて実用的システムを構築した点で優れている/新規性がある。

認知機能測定手法の確立は、認知症の言語による解析という今後ますます研究される分野のテストベッドを構築した点で注目に値する。

これら 3 つの成果はそれぞれ単独の研究としても大きな成果であるが、さらに、これらを総合し、一貫した考察を与えた本研究の意義は極めて高いと考えられる。

なお、最終審査において、章の構成やグラフの記載方法について、軽微な修正やアドバイスをいただいた。これらが反映されたこともすでに確認している。以上をもって、博士(工学)の学位論文としての価値があるものと認める。