

論文内容の要旨

申請者氏名 西山 智弘

近年、自然言語処理(NLP)の進展により、電子カルテ(EHR)やソーシャルメディアなど、非構造化医療データから有用な情報を抽出する技術が実用化の段階に近づいている。本論文では、こうしたデータを臨床現場で実用的に活用するための方法論を探求し、情報の抽出、構造化、適用における課題を解決することを目指す。ヘルスケア分野において正確で包括的なデータの利用は、医療の質と安全性を向上させるために不可欠である。そこで、本研究では非構造化データである医療文書とソーシャルメディアの2種類のテキストからの情報抽出とその利用について実用化の方法を検討する。医療文書については、固有表現抽出(NER)技術を用いた情報抽出と構造化の方法を検討し、従来の構造化データでは得られなかった症状や副作用の可視化を可能にする手法を開発する。ソーシャルメディアデータからの情報抽出については、ソーシャルメディアデータを用いて医薬品ノンコンプライアンスの発言を分析するモデルを構築する。このモデルは転移学習を活用し、限られたデータ環境下でも高い性能を示す。さらに、薬物乱用とデータ不足という2つの課題に取り組む。1つ目の課題については、ソーシャルメディアにおける薬物乱用の実態を分析する新しい分類フレームワークを開発し、乱用の傾向や影響を可視化する。2つ目の課題に対しては、生成モデルを用いたデータ拡張により、学習データ不足がモデル訓練に与える影響を解決する方法を提案する。生成されたデータと実際の医療統計を比較し、補完的な学習データとしての有用性を検証する。本論文では、非構造化データである医療文書とソーシャルメディアからの情報抽出とその利用に焦点を当て、症状や副作用の可視化、ノンコンプライアンスの分析、データ不足の克服といった課題への取り組んだ。情報不足への対策や薬物乱用の可視化といった課題への対応を通じて、医療の質と患者安全の向上に寄与することを目指している。

論文審査結果の要旨

申請者氏名 西山 智弘

令和7年2月7日、審査委員一同は本論文申請者に対し、論文の内容および関連事項についての審査試験を実施した。審査においては、論文の学術的価値、研究の独自性、方法論の妥当性、ならびに論理構成の明確性について詳細に検討を行った。

試験の結果、申請者は審査過程において指摘された問題点について十分な理解を示し、適切かつ的確に回答を行った。また、論文本体に対して求められた修正・加筆も、内容の深化および明確化に寄与する形で実施されており、その修正は学術的にも十分に意義のあるものであった。以上の審査結果を総合的に判断し、当該論文は博士論文としての要件を満たすものであると認める。