



Title	Extracting clinical terms from radiology reports with deep learning
Author(s)	杉本, 賢人
Citation	大阪大学, 2023, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/92054
rights	
Note	やむを得ない事由があると学位審査研究科が承認したため、全文に代えてその内容の要約を公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

論文内容の要旨
Synopsis of Thesis

氏 名 Name	杉本 賢人
論文題名 Title	Extracting clinical terms from radiology reports with deep learning (深層学習を用いた放射線レポートからの臨床用語の抽出)
論文内容の要旨	
[目 的(Purpose)] 放射線レポートには多くの臨床情報が書かれているが、フリーテキストで書かれているため、そこから情報抽出するのが難しい。二次利用に向けて、放射線レポートから臨床用語を抽出することは、最初の重要なステップである。しかし、抽出される用語の種類については一般的なコンセンサスが得られていない。そこで、本論文では、観察・臨床所見・修飾語という3種類の臨床エンティティからなる情報モデルを提案する。さらに、放射線レポートへの適用性を判断するため、最先端の深層学習モデルを用いて臨床用語を抽出し、その結果を比較する。	
[方法ならびに成績(Methods/Results)] 複数の医療従事者がアノテーションした胸部コンピュータ断層撮影 (CT) レポート540件を用いて、モデルの学習と評価を行った。2つの深層学習モデルを比較し、事前学習の効果を検討した。また、モデルの一般化可能性を調べるために、他施設の胸部CTレポートを用いて評価した。院内データセットと他施設データセットを用いた我々の最高性能モデルのF1スコアは、それぞれ95.36%と94.62%であった。	
[総 括(Conclusion)] 我々の情報モデルで定義されたエンティティは、放射線レポートから臨床用語を抽出するのに適しており、また、このモデルは他の施設のデータセットで使用するのに十分な汎化性を持っていることが示唆された。	

論文審査の結果の要旨及び担当者

(申請者氏名) 杉本 賢人				
論文審査担当者	(職)	氏 名		
	主 査	大阪大学教授	武田 理宏	〇
	副 査	大阪大学教授	新谷 康	〇
	副 査	大阪大学教授	服部 聡	〇

論文審査の結果の要旨

電子カルテの普及が進み、リアルワールドデータの利活用が期待されているが、現在の利活用の多くは、レセプト、薬剤、検査結果など構造化されたフォーマットで保存されている一部のデータに限定されている。電子カルテには多くのフリーテキストで書かれた非構造化データが保存されており、これらを二次利用可能にすることは、医療情報学が取り組むべき大きな研究課題のひとつとされている。本論文は、非構造化データの1つであり、重要な臨床情報を多く含んでいる放射線レポートの構造化を目的とし、レポートに書かれた臨床情報を深層学習により正確に取り出すことに取り組んだ論文である。論文では、大阪大学医学部附属病院の胸部CTレポートを使用して、複数の医療従事者によってラベル付けを行い、それを複数の深層学習モデルを用いて実験している。実際の臨床データを用いて評価を行い、モデルのF1 scoreは95.36%と良好な性能であったことから、今後、放射線レポートの構造化を介し、リアルワールドデータの利活用が活性化するための重要な課題を克服している点から、学位に値すると考える。