

放射線レポートに含まれる悪性腫瘍に関する情報の 未認識を検出するプログラムの開発

横田慎一郎^{*1}、土井俊祐^{*1}、福原正和^{*1}、三谷知広^{*1}、永島里美^{*1}、今井健^{*2}、大江和彦^{*1, *2}

^{*1} 東京大学医学部附属病院, ^{*2} 東京大学大学院医学研究科

Developing Application Programs to Detect Unrecognition of Information about Malignant Tumors in Radiology Report

Shinichiroh Yokota^{*1}, Shunsuke Doi^{*1}, Masakazu Fukuhara^{*2},
Tomohiro Mitani^{*2}, Satomi Nagashima^{*1}, Takeshi Imai^{*2}, Kazuhiko Ohe^{*1, *2}

^{*1} The University of Tokyo Hospital, ^{*2} Graduate School of Medicine, The University of Tokyo

抄録:

放射線読影レポートによる適切な情報伝達は診療において重要な位置を占めており、依頼医によるレポートの見落としの予防は国内外を問わず重要な課題である。本研究では、レポート内に読影医が記載した疾患名が、レポート作成以後の期間に患者の診療記録内に出現しない場合、依頼医がその疾患について認識しておらず、すなわちレポート中の重要な所見を認識していない可能性があるとの前提に基づき、1) レポートから疾患名を抽出する機能、2) 抽出した疾患名についてレポート作成以降で一定期間の診療記録内を検索し、それが出現しなかった疾患名に対応する標準病名とその ICD10 コードを出力する機能、3) 処理を毎日バッチで自動実行し結果をメール通知する機能、を開発した。検証結果を踏まえると、レポートの見落とし状況を第三者が確認する際の、膨大とも言える診療記録の点検を支援できる可能性があり、今後実運用にむけたさらなる研究開発を行いたい。

キーワード 医療安全, 電子カルテ, 放射線読影レポート, 読影所見の確認漏れ, 自然言語処理

1. はじめに

米国放射線専門医会や英国王立放射線科専門医会は、放射線読影レポート(以下、レポート)にかかる重要所見に関するコミュニケーションの重要性を説いている。本邦でも厚生労働省、日本医療機能評価機構、日本医学放射線学会が注意喚起等を発出しており、レポートにかかる適切な情報伝達は、国内外で重要な課題である。レポートが主治医に認識されない要因として日本医学放射線学会は、放射線科診断専門医の不足や医療者間のコミュニケーション不足を挙げている。日本は経済協力開発機構加盟各国と比べて放射線検査数が多く、伴ってレポートも多いと推測できることから、松村らの報告書[1]や日本学術会議の提言に記載があるように、情報システムによる支援が欠かせない。一概にレポートの見落としと言っても、依頼医側がレポートの存在に気づいていない、存在に気づいても読んでいない、読んでも読影医の意図を認識していないあるいは誤解している、等の状況がある。したがって、未読

かどうかの量的な点検に加え、より発展的にはレポートや診療記録の内容を質的に考慮した点検が必要である。依頼医が認識する疾患名は全て診療記録内に文字列として出現すると仮定すると、レポート内で読影医が記載した疾患名がその後の診療記録内に出現しない場合、依頼医が当該疾患名を認識しておらず、すなわちレポート見落としの可能性もある。人間が大量のレポートや診療記録の内容を網羅的に点検し、意味的な見落としを検出する事は非現実的であるため、機械的な検出方法の開発が必要である。

2. 目的

読影医がレポートに記載した疾患名のうち、診療記録に出現しない疾患名を検出することで、読影医の意図が依頼医に正しく認識されていないことを検出・通知するプログラムを開発する。

3. 方法

1) システム実装

次の 3 機能を実装した。1) 読影レポートシステムから Web スクレイピングにより CT 検査に関する

レポートを抽出し、専用に構築した文章分類器を用いて現在アクティブな悪性腫瘍に関する疾患名をレポートから抽出。2)電子カルテシステムのレプリケーションデータベース内のレポート作成後14日間分の診療記録に対して、抽出した疾患名と標準病名マスタによる類語を全文検索し、出現しない標準病名とそのICD10コードを出力。3)処理を毎日自動実行し、結果をメール通知。

2) 文章分類器の構築

レポート内には治療済みの病名や否定表現の疾患名も存在するため、不要なものを除外する必要がある。本研究では見落としを警告すべき重要性が高い疾患として悪性腫瘍を選択し、実際のレポート内の悪性腫瘍について記載されている文章を材料として改行と句点で分割し、Bag of words でベクトル化した883文章を用い、Support Vector Machine (SVM)、決定木、ナイーブベイズの3方法と、MeCab並びに万病辞書[2]により、悪性腫瘍に関する標準病名を含む文章が「現在疾患がある事を示す文章」か否かの分類器を構築した。構築した分類器のうち、5分割交差検証により最も良い成績を収めた分類モデルで実装した。

3) 疾患名抽出内容の検証

2020年11月下旬の8日間に作成されたレポート1,194件からプログラムで自動抽出された疾患名について、診療記録を目視確認した。

4. 結果

1) システム実装

Pythonを用いてWindows 10上に実装した。

2) 文章分類器構築の構築

最も良好な成績を示したのはSVMで、5分割交差検証での平均成績は、正解率: 0.917, 陽性的中率: 0.931, 感度: 0.777, F値: 0.847, Receiver Operatorating Characteristic 曲線のArea Under Curve: 0.972、であった。上記の結果によりSVMを用いて実装することとした。

3) 疾患名抽出内容の検証

対象1,194件のうち、悪性腫瘍に関する疾患名が抽出されたレポートは51件あり、このうちレポート作成後14日間の診療記録内に疾患名の記載が無かったケースが7件あった。この7件を目視確認した結果、標準病名マスタの索引語だけ

では拾えない表記揺れを除いても、全く出現しなかった疾患名が2個存在した。なお診療記録内の治療内容からいずれも対処中と判断できた。

5. 考察

診療記録内に表記のない病名を実際に抽出でき、レポートの見落とし状況を第三者が確認するための新規性のある仕組みを構築できた。文章からの病名抽出については荒牧らの「日本語病名抽出システム MedEX/J」がある。手法が違うことから直接的に比較できないが、部分的な比較を試みると、荒牧らによる単語ベースのアプローチ[3]と本研究による文章分類器の成績は比肩している。なお今回の方法では、検査依頼医の真の認識漏れを文字列の出現有無で代用して検出する形であるため、真の陽性的中率を検証できていない点が研究の限界である。また残課題として、メール通知を誰が受信し、誰が誰に確認するのか実運用と照らして検討が必要である。

6. 結語

読影医がレポート内に記載した疾患名のうち、患者の診療記録に出現しない疾患名を検出することで、読影医の意図が依頼医に正しく認識されないことを検出・通知するプログラムを開発した。レポートの見落とし状況を第三者が確認する際の、膨大とも言える診療記録の点検の際を支援できる可能性があり、今後実運用を開始したい。

参 考 文 献

- [1] 松村泰志, 他. 厚生労働科学指定研究「医療安全に資する病院情報システムの機能を普及させるための施策に関する研究」報告書. <https://mhlw-grants.niph.go.jp/niph/searh/NIDD00.do?resrchNum=201821058A>
- [2] 万病辞書 <http://sociocom.jp/~data/2018-manbyo/index.html>
- [3] 荒牧英治, 他. 病名アノテーションが付与された医療テキストコーパスの構築. 自然言語処理. 2018: 25(1); 119-15. 2017: 245; 285-288.