

最終試験の結果の要旨

報告番号	総研第 614 号	学位申請者	小牧 祥太郎
審査委員	主査	武田 泰生	学位 博士 (医学・歯学・学術)
	副査	郡山 千早	副査 吉浦 敬
	副査	於保 孝彦	副査 大脇 哲洋

主査および副査の5名は、令和3年6月1日、学位申請者 小牧 祥太郎 君に面接し、学位申請論文の内容について説明を求めると共に、関連事項について試問を行った。具体的には、以下のような質疑応答がなされ、いずれについても満足すべき回答を得ることができた。

質問1) 看護観察記録において、ある特定の用語に焦点を当てて解析したのか。

(回答) 本研究は、看護観察記録を文章ベクトル化することで、コサイン類似度として過去の誤嚥性肺炎症例の記録との類似性を評価している。そのため、記録中の特定の用語に焦点を当てた解析を行ってはいない。

質問2) 大学病院以外の病院の記録でも同様の解析ができるか。

(回答) 2005年に施行された個人情報保護法を受けて、鹿児島大学病院では記録に関する研修・教育がなされている。よって、大学病院の看護師は標準的な記録の記載方法に関して教育を受けており、記録の質は担保されていると考えている。しかし、すべての医療機関において大学病院と同等の水準で記録されているとは言い難い。そのため、他の医療機関で行なった場合には精度は低下すると考えるが、今後、看護師の記録の標準化が進んでいく事により、同水準の解析が行えると考える。

質問3) 誤嚥性肺炎の診断の正確さはどの様に担保されているか。

(回答) 誤嚥性肺炎と確定するには決め手に欠けるが、今回は、診断名、抗生剤の処方の有無、胸部X線写真、また患者のエピソードを鑑みて、共同著者の医師と確認して誤嚥性肺炎症例と決定した。本研究では、「誤嚥性肺炎」症例の選定作業が最も苦勞した点ではあった。

質問4) 各群の設定(各データセットのデータ数やデータの配置について)はどのような基準で行なったか。

(回答) 対照群は、発症群と同じ診療科と年齢構成の患者を抽出している。なお、各データセットの配置はランダムとしている。今回、予備調査データで機械学習パラメーターの事前調査を行う必要より、予備調査のデータ数は多く設定した。また、今回の疾患はデータ数が少なく、どの程度データが集まるか見込みが立ちにくい状況であった。その為、テストデータセットのデータ数は少なく設定せざるを得ない部分が生じた。

質問5) 4日間の時間軸(1-2日と3-4日など)の中での変化はないのか。

(回答) 本研究では、4日間の記録を1データとして扱っており、発症前の1-2日と3-4日といった時間軸での検証は行っていない。記録は、発症の予兆となる患者の様子となるため、大きな変化はないと予想されるが、時間軸における検証も必要であると考えている。

質問6) テストデータの方が適合性やAUC (Area Under the Curve) が良くなっているのはなぜか。

(回答) サンプルデータが少ないため、効果がやや大きく出たことは考えられるが、予備調査データより著しく成績は向上していないため、予測できる結果の範囲内と思われる。

最終試験の結果の要旨

質問 7) 前回手法 (TF-IDF) と今回手法 (Doc2vec) の差は何か。

(回答) TF-IDF においては、誤嚥性肺炎症例における特徴的な語彙を抽出できたが、例えば、「むせ込み」の用語は抽出できても、その有無までは判定ができない。また、特徴語彙の抽出のみであり、主観性の域を脱しない点がある為、客観的に数値として比較可能である本手法を使用した。

質問 8) 文章中の単語を N 次元空間に配置とはどのような意味か。

(回答) 座標空間を意味し、単語が変化しても意味が大きく変わらないような、意味的関連が近い単語はベクトル (大きさや方向性) が近くなる (例: 「男」と「女」は近く、「男」と「りんご」は遠い関係)。この、単語の空間上の配置を意味し、なお、ベクトルの次元数はこちらでパラメータ値として調整を行う。

質問 9) 看護師の記録者による記録の質に関する変化はないか。

(回答) 若干の記録の差異は考えられるが、教育を受けた看護師が患者状態をシンプルに記載しているため、大きな変化はないと考える。なお、記載者はその症例について過去に記載した職員の記録も確認しており、記載内容はある程度類似して記載されることも考えられる。

質問 10) 前回手法の TF-IDF は今回、単語の選択に使っていないのか。

(回答) 本研究では、パラメータ値である min_count 範囲の設定のためだけに使用した。

質問 11) 濁字・片仮名・英語などの表記による問題はどうか。

(回答) TF-IDF では問題点であったが、本研究で使用した機械学習ツールでは文脈で判断するため、同様な文章として扱われる。そのため、表記揺れの問題は生じにくい。

質問 12) 同じ診療科の中で症例データ (記録) を集めた方が精度は上がるのではないか。

(回答) 単一診療科であれば、発症群と対照群の全体的な記録内容は近似すると考えられ、その中で発症群は誤嚥性肺炎が予測される特徴的な記載となると考えられる。そのため、精度向上は期待されるが、今回は単一診療科のみで十分なデータ数を確保できず、全科よりデータを集めた。

質問 13) 誤嚥性肺炎は慢性期の医療機関でよく扱われるが、大学病院の記録を扱うのであれば、大学病院でよく扱われる疾患やインシデントで行ってみてはどうか。

(回答) 大学病院でもインシデント報告としてしばし開かれる「転倒・転落」などは、本研究手法を用いて発症予兆へ活用可能であると考えられる。

質問 14) 2013-2018 年の記録を扱っているが、時期 (年度) による看護師の記録の書き方の違いの影響はないか。

(回答) 看護観察記録はシンプルな記載をされており、患者の入院中の状況に着目して記載されるという点は変わらないため、違いは少ないと考える。

質問 15) コサイン類似度はどの様に見れば良いか。

(回答) 類似度判定の計算手法であり、0-1 の範囲に正規化される。1 に近いほど類似していると考えられる。

質問 16) データ数を増やした方が良いという中で、コサイン類似度を算出する学習用データ数を今回、上位 8、9 例と制限した方が精度が上がったのはなぜか。また精度向上するために必要なものはあるか。

(回答) 誤嚥性肺炎は下位分類が多いといった特徴があるため、比較を行う記録の誤嚥性肺炎のタイプに近似したデータを抽出する処理を行う必要がある。よって、コサイン類似度算出用の学習用データは全例を使わない方が精度は向上した。学習用データのデータ数は多い方が精度は向上すると思われ、それに伴い、一定の割合で算出する学習用データ数の上限は増加すると考える。

以上の結果から、5名の審査委員は申請者が大学院博士課程修了者としての学力・識見を有しているものと認め、博士 (医学) の学位を与えるに足る資格を有するものと認定した。