

厚生労働科学研究費補助金（地域医療基盤開発推進研究事業）

数理最適化モデルによる小学校区グリッドに基づく多組織連携システムの解析（中尾博之研究代表者）

分担研究報告書

クロノロジー（活動記録表）を利用した業務優先度の分析に関する研究

研究分担者 竹内孔一（大）岡山大学学術研究院 自然科学学域 准教授

研究要旨：大規模災害時における地域の災害医療活動能力の資源の再配分を行うにあたり、時々刻々と発生する業務の優先度を適切に判断することは最適な資源配分を実施する上で重要な課題である。本研究の目的は災害時における活動記録であるクロノロジーの業務に対して記載された内容を基に優先度を付与する最適な数理モデルとその適用方法について最善の手法を明らかにすることである。**方法：**災害時において県保健医療調整本部等で記録されるクロノロジーには、災害時における連絡、要求、指令を含む活動が多岐にわたって記録されている。クロノロジーの内容を災害医療活動の周辺で発生する医療活動に関連する文書と捉えて、優先度を重要度と緊急度の観点からクロノロジーに対して人手で分析する。分析データに対して複数の最適化モデル（識別モデル）を適用して、優先度を推定する最も適切なモデルと方法を明らかにする。**結果：**（1）人手による分析の結果、クロノロジーに含まれる優先度の高いものは全体の約4%、中程度のものは約26%であった。（2）優先度の高い業務の識別にはニューラルネットワークを用いた手法が高い精度を示した。特に、ニューラルネットワークは他のモデルに比べて再現率が高いことが明らかになった。**まとめ：**本研究結果は災害時に記録されるクロノロジーを利用して記載内容から業務の優先度を判定するシステム構築に向けてニューラルネットワークを利用した手法が有効であることを示している。

研究協力者

山崎 瑠：岡山大学大学院博士前期課程2年

石澤 哉子：岡山大学大学院非常勤研究員

齋藤 由美：岡山大学大学院非常勤研究員

A. 研究目的

背景：大規模災害時における地域の災害医療活動能力の資源の再配分を行うにあたり、発生する業務の優先度を適切に判断することは最適な資源配分を実施する上で重要な課題である。災害時において発生する業務はその断片が保健

医療調整本部や病院等でクロノロジー¹⁾に記録されている。クロノロジーから災害時には業務の引き継ぎに関する連絡事項など優先度の低い業務から、負傷者などの患者に関する優先度の高い業務が発生していることがわかる。現在では、こうした業務に対して、災害対策本部で人により優先度を判定して医療活動の資源の配分を決定している。被害状況と機能している医療機関の状況を勘案した判断は、デジタル化した情報を最大限に利用することで最適化モデルが適用可能になり、より現場の状況に即した判断

が期待される。

本研究の目的は現段階でデジタル化されているクロノロジーを基に、優先度判定するもっとも適切な最適化モデル（識別モデル）とその適用手法を明らかにすることで、業務の優先度を判断するシステム構築に向けたモジュールの構成について考察することである。

B. 研究方法

クロノロジーに対して人手の作業により優先度を重要度と緊急度の観点から3段階で分析したデータを作成する。表1に優先度の3段階の説明を示す。これによりどの程度優先度が高い業務が発生しているのかを明らかにする。次に分析データに対して、複数の識別モデルを適用して推定精度を測定する。これにより、識別モデルがどの程度有効に働くかを明らかにする。

表1. 優先度を判定する基準

分類	判断する基準
高	重要かつ緊急度が高い内容 人の生命に関わる内容は重要度が高い
中	重要または緊急度のどちらか一方が高い内容
低	上記のどちらでもない内容

識別モデルとして SVM²⁾、XGBoost³⁾、4層ニューラルネットワーク⁴⁾を適用する。識別モデルのプログラムとして、SVMはpythonのscikit-learnで配信されている実装を利用する。線形カーネル関数を利用し、パラメータはデフォルトの設定を利用する。XGBoostもpython上の実装を利用する。クラス分類モデルを利用し、学習回数は50回とする。4層ニューラルネットワークは入力層が単語の語彙数分のユニッ

トを設定して、以降、300、100、3個のユニットを設置する。学習回数は40回とする。

識別モデルに対する入力として文書を形態素解析して単語のベクトルに変換する。具体的には bag-of-words 形式を利用する⁵⁾。形態素解析には固有表現が多く登録されている Neologd 辞書⁶⁾を用いた MeCab⁷⁾を適用する。

識別モデルの学習には学習データが必要である。優先度を付与した分析データの8割を学習データに設定して2割をテストデータとして評価に利用する。

(倫理面への配慮)

本研究では、被災者の個人情報を利用していない。ただし、研究を遂行する上で、倫理的側面を十分に配慮しながら実施した。なお、本研究分担者・協力者は、研究分担者と同じ組織に所属している。

C. 研究結果

1. クロノロジーの優先度分析

クロノロジーの優先度を2名作業員で分析した結果を表3に示す。まず作業員間の一致度は0.805であった。作業員Aの方が、作業員Bに比べて、優先度を高く評価する傾向にあった。

表3. 作業員2名による優先度の付与結果

作業員	作業員 B				
	高	中	低	合計	
A	高	21	39	3	63
	中	15	184	199	398
	低	6	114	1332	1452
	合計	42	337	1537	1913

全体は1913件あり、2名の作業員のうち、両

方とも優先度で「高」を付与した場合は 21 件であり、全体の 0.2%であった。

次に、作業者の分析結果を集約する。2 名の作業者が付与した優先度を平均して最終判定として、分析データを作成した。これは識別モデルに対して、優先度が少しでも高いと判定できるものを手かがりとして学習させるためである。

分析データを 8 対 2 に分割して、学習データとテストデータを作成する。学習データとテストデータの分布を表 4 に示す。

表 4 の集約したデータの内訳を見ると、優先度「高」は 4% (75 件)、優先度「中」は 26.5% (506 件) であった。よって、優先度を付与することは多くの優先度が低いクロノロジーの中から高いもの取り出すことを意味する。そこで、識別モデルの評価の際には、優先度が「中」以上の識別能力を重視してモデルを考察する。

表 4. 学習データとテストデータの分布

クラス	学習データ	テストデータ	全体
高	60	15	75
中	413	93	506
低	1057	275	1332
合計	1530	383	1913

2. 識別モデルによる優先度推定実験の結果

各識別モデルを用いてテストデータの優先度を判別した際の分類精度を表 5 に示す。精度は下記の式で定義する。

$$\text{精度} = \frac{\text{優先度クラス的一致数}}{\text{全件数}}$$

また、表の括弧内の数字は件数を示している。

表 5 の精度からどの識別モデルも近い値であることがわかるが、その中で XGBoost が最も高

い精度を示しており、精度は約 0.8 であった。

一方、クロノロジーの中で重視するものは優先度が高いものである。そこで、優先度「高」および優先度「中」のクラスに対する識別能力を表 6 と表 7 にそれぞれ示す。ここで評価尺度として適合率と再現率、F 値を利用する。それぞれの評価式は下記のように定義する。

$$\text{適合率} = \frac{\text{クラスが一致した件数}}{\text{モデルが対象クラスと判定した件数}}$$

$$\text{再現率} = \frac{\text{クラスが一致した件数}}{\text{データ内の対象クラスの件数}}$$

$$\text{F 値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

表 5. 各モデルの精度

モデル	精度
SVM	0.762 (292/383)
XGBoost	0.799 (306/383)
4-layer	0.778 (298/383)

表 6. 「高」クラスの適合率、再現率、F 値

モデル	適合率	再現率	F 値
SVM	0.357(5/14)	0.333(5/15)	0.345
XGBoost	0.500(1/2)	0.07(1/15)	0.117
4-layer	0.500(6/12)	0.40(6/15)	0.444

表 7. 「中」クラスの適合率、再現率、F 値

モデル	適合率	再現率	F 値
SVM	0.564(44/78)	0.473(44/93)	0.515
XGBoost	0.657(44/67)	0.473(44/93)	0.550
4-layer	0.568(53/95)	0.581(54/93)	0.574

表 6 の結果では「高」クラスの識別では 4 層ニューラルネットワークが F 値で高い値を示し

た。表 5 で高い値を示した XGBoost は再現率が低く、優先度の高い事例を識別できていない場合が多いことがわかる。また、再現率で比較すると 4 層ニューラルネットワークが最も高い値を示した。このことから「高」クラスのカテゴリでは 4 層ニューラルネットワークが良いことがわかる。

次に、表 7 の「中」カテゴリーのカテゴリ分けの結果を見ると、SVM と XGBoost は再現率に関して全く同じ値であった。また適合率に関しては XGBoost が高い値を示した。しかしながら、再現率、および F 値は 4 層ニューラルネットワークが最も高い値を示した。本研究の目的である優先度判定システムはシステムが推定した優先度を人が最終的に判断して利用することを仮定している。このような枠組から再現率の高いモデルの方が利用価値が高いと考えられる。よって実験の結果から、学習データを仮定した上で、優先度を判定するモデルとして 4 層ニューラルネットワークが候補になることが明らかになった。また、その際の参考となる識別能力は F 値で 0.57 であることが明らかになった。

D. 考察

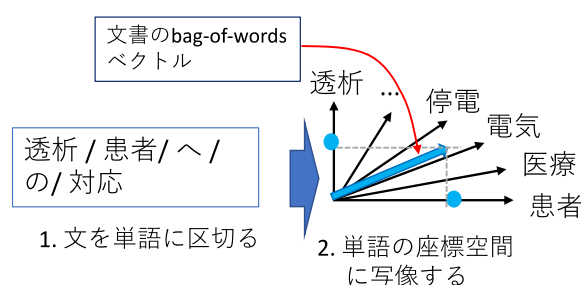
推定モデルの開発には常に精度を改善する手法を検討することが重要である。そこで考察では優先度推定モデルにおけるベクトル化の拡張可能性について議論し、精度向上の余地について議論する。また実際の識別結果を取り上げて、改善に必要な要素について論じる。

1. ベクトル化手法の拡張性

推定モデルの拡張を検討するために、文書をベクトルに変換する単語区切りに対する分析と近年のベクトル化手法について論じる。

1) 単語区切りの具体例と拡張可能性

クロノロジーの文書に対して、識別モデルで分類する際に、文書をベクトル化する必要がある。上述の実験では、bag-of-words 法を利用した (図 1)。Bag-of-words 法を最大限に有効活用するには単語の切り方が重要である。単語区切りは形態素解析器を適用するが、本実験では Neologd 辞書を利用して、病名や固有名詞を可



能な限り正しく区切る手法を適用した。

図 1. 文書をベクトル化する bag-of-words

表 8 に、クロノロジー内の文を形態素解析した際の、正しい区切りの例と誤りの例を提示する。ここで、スラッシュ記号「/」は形態素解析器が分割した単語の区切りを表す。

表 8. 単語区切りの例

区切りが正しい例	正しくない例
「1 人 / 脳梗塞」「熱中症」「抗凝固薬」「セルシン」	「二 / 万 / 橋 / 救護所」「リハ / センター」

表 8 の正しい例では「抗凝固剤」など医薬品名が辞書に登録されていて、一つの単語として取り出されている。一方で、正しくない例では、「二万橋」などの地名や、「リハセンター」などの略語 (正しくは「リハビリセンター」) を正しく分割することができていない。地名は辞書の

拡充で対応できることが考えられるが、略語に関しては現段階では対応する明確な方法はない。クロノロジーの性質として、災害対策本部では多くの情報が発生するため、記述の際に略語が使われるのは自然である。このような略語の処理は今後の改善における一つの要素であると考えられる。

2) 近年の特徴ベクトル化手法

文書をベクトル化する手法として近年、深層学習モデルを適用した手法が提案されている。それらの手法では、日本語の Wikipedia など大量な日本語テキストデータさえあれば、そのテキストデータから日本語の文字列内の連続性に内在する意味的な構造をニューラルネットワークの中に埋め込むことができる事前学習モデルが提案されている⁸⁾⁹⁾¹⁰⁾。事前学習モデルを適用することで、多種の自然言語処理タスクの精度が改善されたことが報告されている⁸⁾。一方で、事前学習モデルは設定すべきパラメータが多く、単純な適用では精度が得られない例も報告されている¹¹⁾¹²⁾。本報告では、基本的な識別モデルの性能を比較するために bag-of-words に注目したが、こうした事前学習モデルの適用は、改善手法の一つとして候補に挙げられる。

2. 優先度識別結果に対する分析

4 層ニューラルネットワークモデルを利用した優先度識別モデルに置いて、推定結果が正しい場合、誤っている場合の例を表 9 に示す。

表 9 において、例 1 と 2 はモデルが正しく優先度を識別したが、例 3 では優先度を 1 段階誤って識別している。上述の表 8 に示したように、単語区切りにおいて、「脳梗塞」「抗凝固薬」が正しく区切られているにもかかわらず、優先度

の判定が低く識別されている。

これは分類モデルとして学習データに「抗凝固薬」と「高」クラスを結び付ける手掛かりとなる表現が不足していたために発生したと考えられる。学習データが不足していることが第一の原因であるため、この改善には学習データを増やすことが上げられる。しかしながら学習データに無い事例を分析することは常に発生するため、補完する機能を追加することが重要である。この機能について、上述のように、近年の事前学習モデルは言語の特徴を取り込んでおり、不足している表現を補完して正しく識別することが期待できる。よって、精度向上に向けて事前学習モデルを取り込むことが重要であると考えられる。

表 9. 4 層ニューラルネットワークの識別結果

	推定	正解	文書
1	高	高	病院協会より周辺病院 (50 病院) に XX 病院の入院患者受入要請
2	中	中	YY に患者 20 名を搬送中
3	中	高	ZZ 周辺薬足りない人 1 人 脳梗塞 抗凝固薬 3 日間で服できていない

文書は一部伏せ字に置換している。

E. 結論

災害時において記録されるクロノロジーを利用して記載内容における業務の優先度を推定するシステムの構築を目標として、クロノロジーの分析および複数の識別モデルを適用した優先度推定実験を行い、適切な適用手法について考察した。その結果、(1) クロノロジーに記載さ

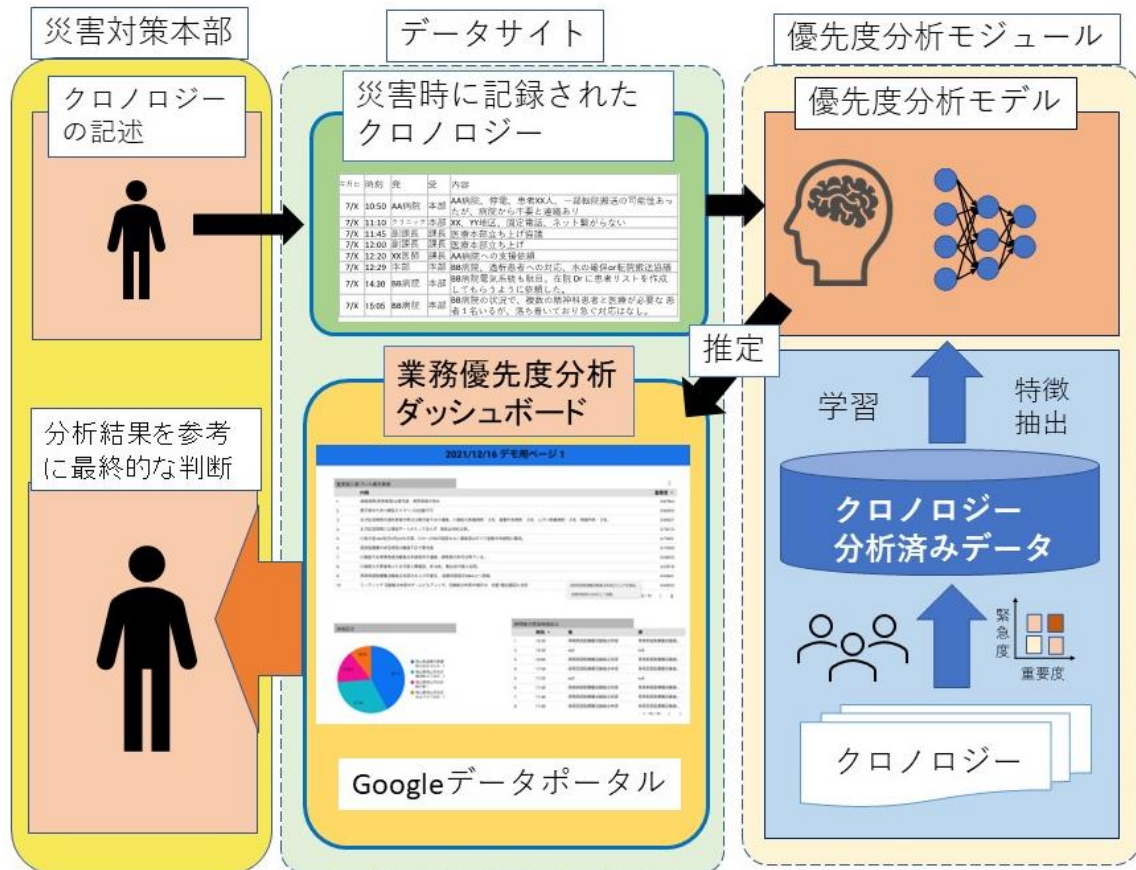
れている業務のうち、優先度が高いものは4%、中程度のものまで含めても3割程度と少なく、

(2) 優先度を推定するモデルとしてニューラルネットワークモデルが効果的であることが実験的に明らかになった。この結果は、災害時に優先業務を決定するシステムの構成要素として、クロノロジーを学習したニューラルネットワークモデルによって優先度を計算するモジュールが適用可能であることを示唆している。図2に優先度計算結果を可視化したデモシステムを示す。Googleのポータルサイトを利用することで、関係者が容易に結果を確認することができる。



図2. クロノロジーを利用した業務優先度の推定結果の可視化デモシステム

優先業務順位決定の概念図



F. 健康危険情報 なし

G. 研究発表

1. 論文発表 なし

2. 学会発表

竹内孔一、山崎瑤、渡邊暁洋、平山隆浩、中尾博之. 災害医療におけるクロノロジーの分析. 電子情報通信学会信学技報. 121(415) NLC2021-31, 19-23, 2022.

H. 知的財産権の出願・登録状況 なし

引用文献

1. 岡山県医師会、災害医療深救護マニュアル、2020.
2. 麻生英樹、津田宏治、村田昇. パターン認識と学習の統計学. 岩波書店. 2003.
3. Chen T, and Guestrin C, XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016.
4. 岡谷貴之. 深層学習. 講談社. 2015.
5. 高村大也. 言語処理のための機械学習入門. コロナ社. 2010.
6. 佐藤敏紀、橋本泰一、奥村学. 単語分かち書き辞書 mecab-ipadic-neologd の実装と情報検索における効果的な使用方法の検討. 言語処理学会第 23 回年次大会、pp.875-878. 2017.
7. Kudo T. MeCab: Yet Another Part-of-Speech and Morphological Analyzer. <https://taku910.github.io/mecab/>
8. Devlin J, Chang M-W, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT, pp. 4171-4186, 2019.
9. Brown T, Mann B, Ryder N, et al. Language Models are Few-Shot Learners. Proceedings of NeurIPS, 2020.
10. Raffel C, Shazeer N, Roberts A, et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. Journal of Machine Learning Research, 21(140):1-67, 2020.
11. 竹内孔一、大野雅幸、泉仁宏太、田口雅弘、稲田佳彦、飯塚誠也、阿保達彦、上田均. 研究利用可能な小論文データに基づく参照文書を利用した小論文採点手法の開発、情報処理学会論文誌. 62(9), 1586-1604, 2021.
12. Mayfield E and Black A W. Should You Fine-Tune BERT for Automated Essay Scoring? Proceedings of the Fifteenth Workshop on Innovative Use of NLP for Building Educational Applications, 151-162, 2020.