

保健師活動必要度判定アルゴリズムを搭載した ICT ツールの
保健師人材育成への活用に関する実証研究

分担研究者 三浦貴大 国立研究開発法人産業技術総合研究所人間拡張研究センター
共創場デザイン研究チーム 上級主任研究員

研究代表者 吉岡京子 東京大学大学院医学系研究科健康科学・看護学専攻 准教授

分担研究者 塩見美抄 京都大学大学院医学研究科人間健康科学系専攻 准教授

分担研究者 細谷紀子 千葉県立保健医療大学健康科学部看護学科 准教授

分担研究者 佐藤美樹 国立保健医療科学院生涯健康研究部 上席主任研究官

分担研究者 藤井仁 目白大学看護学部看護学科 教授

研究要旨

本研究では、前年度に開発した保健師活動内容の入力用 ICT 試作版ツールで得られた入力結果を基に、保健師の判断を模擬する分類モデルの構築・評価を行った。この ICT 試作版ツールは、統括保健師が人材育成を推進する上で必要なデータ収集を円滑に行えるよう、各保健師が日々の活動内容（活動種別、所要時間、アセスメント結果等）を入力可能なものである。

学習用データには、アセスメント項目 21 点に加え、各種アウトカムとして、アセスメント結果より得られた必要度判定結果について保健師が妥当性評価した結果や、保健師の主観に基づく支援必要度に加え、在宅生活を継続・死亡・入院・入所・虐待・自傷・他害の状況が含まれる。本データはデータ拡張によりデータの不均衡さを抑制した。学習には、ランダムフォレストなど 6 種を用いた。

結果より、不均衡さを抑制したデータで学習させた分類モデルが高い性能を示した（Kappa 係数と F-measure）。さらに、このような高い性能だった分類モデルによる特徴量重要度においては、学習方法が異なっても、重み付けが上位であるアセスメント項目が共通する点を確認できた。これら項目は、現場経験が豊富な保健師にとっても違和感がないものであった。一連の結果は、データの不均衡を抑制して拡張したデータを用いて学習させたモデルによって、熟練保健師に近い判断を行える点が示唆された。このモデルの活用方法は今後の課題である。

研究協力者

松本 珠美 公益社団法人日本看護協会
常任理事

山下久美 兵庫県北播磨県民局加東健康福祉
事務所 参事

松下美穂子 杉並区子ども家庭部兼保健福祉
部地域保健調整担当課長事務取扱

宇井 教恵 千葉県横芝光町健康子ども課
健康づくり班 主幹

茂木 りほ 国立保健医療科学院生涯健康
研究部 主任研究官

本田 千可子 東京大学大学院医学系研究科
健康科学・看護学専攻地域看護学分野 助教

松本 博成 東京大学大学院医学系研究科
健康科学・看護学専攻地域看護学分野 助教

平 和也 京都大学大学院医学研究科
人間健康科学系専攻地域健康創造看護学
助教

角川 由香 東京大学大学院医学系研究科
健康科学・看護学専攻高齢者在宅長期ケア
看護学分野 助教

A. 研究目的

人口急減・超高齢化が進展している我が国では、様々な分野における情報通信技術 (ICT; Information and communication technology) の活用とデジタルトランスフォーメーション (DX; Digital transformation) の必要性が謳われている^{1,2)}。また、新型コロナウイルス感染症 (COVID-19; Coronavirus disease 2019) の流行を契機に迅速な感染防止対策の必要性が再考され、医療業務やそのマネジメントの効率化が求められたことで、医療・看護・介護現場における ICT 導入が急速に進められてきた³⁾。

一方で、この ICT 導入は限定的にしか展開されており、都道府県や市区町村といった地方自治体における保健師活動においては未だ発展途上である。これは各保健師の活動が、各々の暗黙知に基づいて遂行されざるを得ない他⁴⁾、自治体ごとの特性の違いに伴う業務の独自性が生じている可能性がある。このような保健師業務における自治体間の共通性・独自性を明らかにできれば、ICT 導入を円滑に行えるため、データ活用による業務効率化や人材育成の進展が出来ると考えられる。このため、前年度において、各保健師が日々の活動内容 (活動種別、所要時間、アセスメント結果等) を入力可能な ICT 試作版ツールを開発した⁵⁾。その上で、パイロット調査を通

して、ユーザビリティ評価を行うとともに、統括保健師が人材育成を推進する上で必要なデータ収集を実施した。

一方で、大規模データを基にした機械学習モデルを構築の上で、人の作業を支援する研究例が情報学の分野で多く報告されている。本研究においても蓄積されたデータを用いることで、保健師育成の支援を行うためのモデル構築が出来ると考えられる。

そこで本研究の目的を、保健師が入力したアセスメント項目と支援対象者の状況を基に、支援対象者の状況推定を行うモデル構築・評価を行うことと設定した。

B. 研究方法

1) データセット

44,163 件の保健師活動に関する入力データである。本データは、前述した ICT 試作版ツールにより取得された。この中から、次のアセスメント項目 21 点に加え、各種アウトカムを対応付けて学習用モデルとした。

- アセスメント項目 (該当/非該当/不明の 3 段階、不明と記載されたものは欠測値として扱った):
 - 1. 急激病状悪化, 2. 医学対応不十分, 3. 治療体制整備困難, 4. 退院後支援, 5. 意思疎通困難, 6. メンタルヘルス課題, 7. 支援拒否, 8. 問題対処能力課題, 9. サービス制度利用, 10. 日常ケア課題, 11. 家族関係課題, 12. 家族問題対処課題, 13. 家族相談不足, 14. 多問題世帯, 15. 生活療養環境不適切, 16. 経済困窮, 17. 虐待, 18. ハイリスク支援, 19. 他機関連絡, 20. 警消連携要求, 21. 近隣トラブル
- アウトカム:
 - アセスメント結果より得られた必要度判定結果を、保健師による妥当性評価した結果 (妥当/高すぎる/低すぎるの 3 段階)

表 1 本研究で用いるデータセット。Raw data の他、データ拡張後のデータ数を掲載する。

必要度判定結果の妥当性評価結果					保健師の主観に基づく支援必要度															
	1	2	3	合計	1 (不必要)	2	3	4	5	6	7	8 (必要)	合計							
Raw data	6697	66	390	7153	389	674	618	365	3402	547	559	346	6900							
w/ SMOTE (1)	6697	6697	6697	20091																
w/ SMOTE (2)	6697	3348	3348	13393																
SCUT	2384	2384	2384	7152	862	862	862	862	862	862	862	862	862	6900						

在宅生活を継続					死亡					入院				
	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計		
Raw data	294	3272	724	4290	3562	4	724	4290	3457	109	724	4290		
w/ SMOTE (1)	294	3272	3272	6838	3562	3562	3562	10686	3457	3457	3457	10371		
w/ SMOTE (2)	294	1636	1636	3566	3562	1781	1781	7124	3457	1728	1728	6913		
SCUT	1430	1430	1430	4290	1430	1430	1430	4290	1430	1430	1430	4290		

入所					虐待					自傷				
	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計		
Raw data	3514	52	724	4290	3425	141	724	4290	3546	20	724	4290		
w/ SMOTE (1)	3514	3514	3514	10542	3425	3425	3425	10275	3546	3546	3546	10638		
w/ SMOTE (2)	3514	1757	1757	7028	3425	1712	1712	6849	3546	1773	1773	7092		
SCUT	1430	1430	1430	4290	1430	1430	1430	4290	1430	1430	1430	4290		

他害					Accuracy										
	0 (非該当)	1 (該当)	8 (期間中に支援終了)	合計	必要度判定結果の妥当性評価結果	保健師の主観に基づく支援必要度	在宅生活を継続	死亡	入院	入所	虐待	自傷	他害		
Raw data	3546	20	724	4290	0.959	0.619	0.824	0.856	0.849	0.850	0.849	0.854	0.854		
w/ SMOTE (1)	3546	3546	3546	10638	0.929	-	0.592	0.890	0.884	0.876	0.867	0.883	0.888		
w/ SMOTE (2)	3546	1773	1773	7092	0.924	-	0.473	0.829	0.846	0.838	0.830	0.851	0.856		
SCUT	1430	1430	1430	4290	0.928	0.586	0.666	0.877	0.860	0.873	0.834	0.871	0.872		

表 2 構築したモデルの性能 (ランダムフォレストの場合)

Kappa													
	必要度判定結果の妥当性評価結果	保健師の主観に基づく支援必要度	在宅生活を継続	死亡	入院	入所	虐待	自傷	他害				
Raw data	0.557	0.342	0.395	0.259	0.346	0.279	0.398	0.294	0.294				
w/ SMOTE (1)	0.893	-	0.219	0.836	0.826	0.813	0.800	0.824	0.832				
w/ SMOTE (2)	0.878	-	0.029	0.709	0.741	0.726	0.709	0.749	0.758				
SCUT	0.891	0.527	0.497	0.816	0.789	0.809	0.750	0.807	0.809				

F-measure													
	必要度判定結果の妥当性評価結果	保健師の主観に基づく支援必要度	在宅生活を継続	死亡	入院	入所	虐待	自傷	他害				
Raw data	0.646	0.483	0.641	0.612	0.674	0.704	0.687	0.745	0.745				
w/ SMOTE (1)	0.929	-	0.583	0.889	0.884	0.876	0.868	0.882	0.887				
w/ SMOTE (2)	0.923	-	0.472	0.789	0.823	0.811	0.799	0.824	0.832				
SCUT	0.927	0.587	0.690	0.875	0.860	0.875	0.838	0.870	0.871				

- 保健師の主観に基づく支援必要度(1-8 の 8 段階)
- 在宅生活を継続, 死亡, 入院, 入所, 虐待, 自傷, 他害 (該当/非該当/期間中に終了の 3 段階)

これらのデータへの前処理として、アセスメント項目 21 点と、アウトカムから 1 項目を選定の上、欠損値除去した後、データ拡張を行ったもの・行わなかったものを用意した。データ拡張は、SMOTE(Synthetic Minority Over-sampling TEchnique)⁶⁾ と SCUT (SMOTE and Cluster-based Undersampling Technique)⁷⁾ を用いて、不均衡データを可能な限り、均衡データになるようにした。データ点数を表 1 に示す。

ここではデータ拡張を行わなかったものを Raw data, データ拡張を行ったものを、SMOTE(1), SMOTE(2), SCUT のように表記した。なお、SMOTE(1), SMOTE(2)はターゲットとする「該当」などのデータ量を増やす戦略を取った一方で、SCUT はデータ量の balan-

スを取る戦略を取るために、時に最も多い反応を減らすようデータ量を操作した。なお、このようなデータ拡張は外傷患者における院内死亡率の予測の研究論文など、頻発しない医学的事象の予測にも使われてきている⁸⁾。

2) モデルの構築・評価

前述したデータセットを学習用:検証用 = 8 : 2 と分けた後、10 分割の交差検証, 3 回繰り返すよう設定した。また、モデル選定基準を leave-one-out 交差検証とした。

本研究では分類モデルを構築する。学習方

法は、ランダムフォレスト, XGBoost (eXtreme Gradient Boosting), XGBoost DART (Dropouts meet multiple Additive Regression Trees)⁹⁾, ニューラルネット, ElasticNet, SVM (Support Vector Machine) の 6 種類とした。

構築されたモデルの評価は、Accuracy, Kappa 係数に加え、F-measure を用いた。また、特徴量重要度 (Feature importance) についても算出し、各モデルがどのアセスメント項目に重みづけて分類を行ったのかを調べた。一連のモデル構築・評価は R (ver. 4.4.2) を用いて実施し、主に caret パッケージ¹⁰⁾ (ver. 6.0.94) を利用した。

(倫理面への配慮)

本研究で得られたデータは、東京大学大学院医学系研究科・医学部倫理委員会の承認を得て取得された (審査番号 2022114NI-(2)、承認日 2022 年 11 月 10 日、審査番号 2023056NI-(1) 承認日 2024 年 6 月 6 日)。

C. 研究結果

1) モデル評価の結果

表 2 に構築した分類モデルにおける性能評価結果の例として、ランダムフォレストで学習した場合の結果を示す。なお、他の学習方法を用いた場合でも、同様の傾向である。正答率においては、Raw data を用いて学習した場合の必要度判定結果・保健師の主観に基づく支援必要度・在宅生活を継続の項目では他のデータセットを用いた場合よりも良かった。死亡・入院・入所・虐待・自傷・他害の項目では SMOTE を用いてデータ拡張した場合の方が良かった。一方で、クラスごとの正答の偏り具合を表す Kappa 係数や、適合率と再現率のバランスを取った指標である F-measure においては、概して Raw data を用いて学習した場合よりも、データ拡張を用いた場合の方が良かった。このとき、必要支援度・在宅生活を継続の 2 項目を除き、最良モデルにおける F-measure において 0.8 以上の性能を実現できた。

2) 特徴量の重要度評価の結果

図 1~2 にランダムフォレスト・XGBoost を用いて分類モデルを構築した場合における、判定妥当性の評価結果・死亡の場合の特徴量重要度をそれぞれ示す。図 1, 2 の両者ともに、Raw data を用いて学習した場合 (左図)、拡張したデータを用いて学習した場合 (中央図) で、分類の手がかりとするアセスメント項目が異なっており、その重み付けの傾向も大きく違っていた。一方で、データ拡張を行った後で分類モデルを構築した場合には、ランダムフォレスト、XGBoost とともに上位/下位に来る指標は共通していた。特に図 1 においては、18.ハイリスク事例の継続支援である、8.本人の問題認識・問題対処能力に課題がある、11.家族の関係性に課題がある、9.何らかのサービス・制度を利用しているの 4 項目が上位 5 項目に入る点が共通していた。図 2 においても、18.ハイリスクの継続支援である、12.家族の問題認識・問題対処能力に課題がある、4.退院直後の支援の 3 項目が上位 5 項目に入る点が共通していた。このようにデータ拡張された後で学習が行われた場合、Kappa 係数や F-measure が高性能のモデルにおいては、他のアウトカムにおいても同様の傾向が確認された。

D. 考察

1) モデル評価の結果

表 2 において、Raw data を用いた場合は、データの不均衡さに伴って最も高い頻度で現れるデータに偏って分類したため、総合的な正答率は高かったものの、Kappa 係数や F-measure が低くなったと考えられる。実際に、Raw data において該当数が少なかった死亡・入所・自傷・互いにおいて、顕著に Kappa 係数、F-measure が低くなっていた。一方で、このようにデータの不均衡さが際立っている場合ほど、データ拡張による Kappa 係数、F-measure の向上が確認できた。特に F-measure

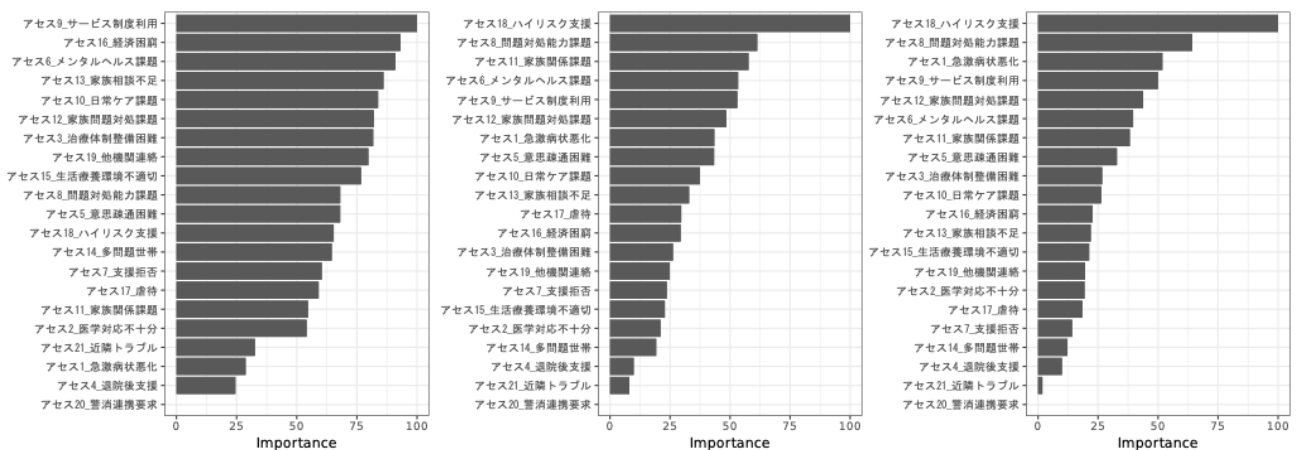


図 1.ランダムフォレスト(左・中図)および XGBoost(右図)を用いて判定妥当性の評価結果の分類モデルを構築した場合における特徴量重要度 (左: Raw data を用いた場合, 中央・右: SMOTE を用いた場合(F-measure において最良ないしはそれに準ずるモデルのとき))

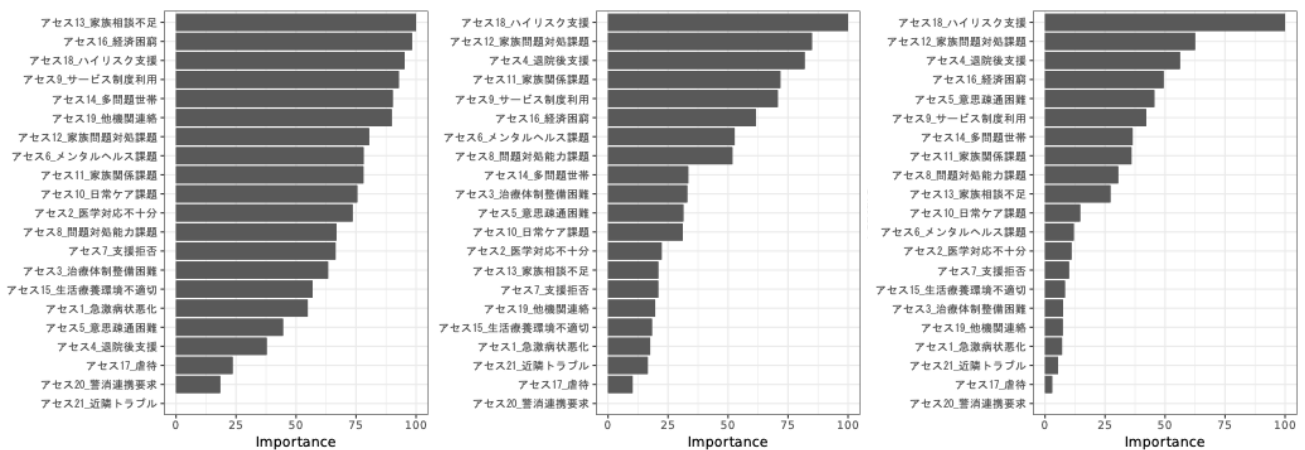


図 2.ランダムフォレスト(左・中図)および XGBoost(右図)を用いて死亡に関する評価結果の分類モデルを構築した場合における特徴量重要度 (左: Raw data を用いた場合, 中央・右: SMOTE を用いた場合(F-measure において最良ないしはそれに準ずるモデルのとき))

の値において、最良モデルでは 0.8 以上の性能を実現できた。一方で、必要支援度・在宅生活を継続の 2 項目を除き実現できなかった項目においては、今後さらなる検討を行う予定である。特に、本研究ではアセスメント項目 21 点のみを用いて分類モデルを構築したため、今後はさらに学習用パラメータを増やすなどして、性能向上を図っていく。

2) 特徴量の重要度評価の結果

Kappa 係数や F-measure が高性能のモデルにおいて、重要度が上位に来るアセスメント項目

は共通した。この点について、経験豊富な保健師に確認した所、実際にこれらの項目をもとに現場での判断を行う可能性がある点が述べられ、違和感がない結果である、との反応を得られた。よって、不均衡なデータであっても、少量ケースをデータ拡張して学習させることで、構築より現場経験のある保健師に近い判断をさせられる可能性がある。

なお、本研究でアウトカムとした状況は、保健師が実際に経験する頻度が稀なものである。現場の保健師は、そのような遭遇しづらい状況の経験をうまく重みづけて、現場のアセスメン

トを効率よく行っていることが示唆される。また、各保健師ともに考え方などは異なっていると考えられるが、本研究でモデル構築に使う学習方法が異なっても多量の均衡データを学習させた場合に類似した特徴量重要度の傾向を示したことから、長きに渡る経験からの判断基準が類似してくる可能性がある点を模擬できた可能性がある。

一方で、遭遇しづらい状況であるほど、経験の浅い保健師におけるアセスメントの学習支援が行いにくい状況とも捉えられる。本研究で得た特徴量重要度に関する知見は、保健師活動の現場支援や現場外での学習支援の方策についての検討に繋がり得るものと考えられる。

E. 結論

保健師による ICT ツールへの入力結果に基づき、アセスメント項目に基づく各種状況の分類モデルを構築・評価した。データ拡張によって、発生頻度が必ずしも高くない事象の分類において、一部項目を除いて、F-measure で 0.8 以上の性能を達成できた。一方で、この分類モデルにおける特徴量重要度を分析した結果、学習手法によらず上位に来るアセスメント項目に共通性が確認できた。

今後は、本研究で構築した分類モデルのさらなる高性能化を図る他、その活用方策についても探っていきたい。

引用文献

- 1) 厚生労働省. 医療 DX 令和ビジョン 2030. https://www.mhlw.go.jp/stf/shingi/other-ise_i_210261_00003.html (2025/4/30 accessed)
- 2) 総務省. 自治体デジタル・トランスフォーメーション (DX) 推進計画 令和 4 年 9 月 2 日. https://www.soumu.go.jp/main_content/000835260.pdf (2024/4/30 accessed)
- 3) 首相官邸 新型コロナウイルス感染症対策分科会. 科学と ICT を用いた対策の提言—多くの国民にワクチン接種が行き渡るまで

- に—令和 3 年 6 月 16 日. https://www.cas.go.jp/jp/seisaku/ful/taisakusuisin/bunkakai/dai4/ict_teigen.pdf (2024/4/30 accessed)
- 4) Yoshioka-Maeda, K., Murashima, S., Asahara, K. Tacit knowledge of public health nurses in identifying community health problems and need for new services: A case study. *International Journal of Nursing Studies*. 43(7): 819-826, 2006.
- 5) Yoshioka-Maeda, K., Matsumoto, H., Honda, C., Taira, K., Hosoya, N., Sato, M., Iwasaki-Motegi, R., Sumikawa, Y., Fujii, H., Miura, T. & Shiomi, M., Development of the Essential Individual Care Needs Assessment Tool for Public Health Nurses. *Public Health Nursing*., 42(3):1216-1225, 2025.
- 6) Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P., SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321-357, 2002.
- 7) Agrawal, A., Viktor, H. L., & Paquet, E., SCUT: Multi-class imbalanced data classification using SMOTE and cluster-based undersampling. In 2015 7Th international joint conference on knowledge discovery, knowledge engineering and knowledge management (IC3k), IEEE, 1:226-234, 2015.
- 8) Hassanzadeh, R., Farhadian, M., & Rafieemehr, H., Hospital mortality prediction in traumatic injuries patients: comparing different SMOTE-based machine learning algorithms. *BMC medical research methodology*, 23(1):101, 2023.
- 9) Vinayak, R. K., & Gilad-Bachrach, R., Dart: Dropouts meet multiple additive regression trees. In *Artificial Intelligence and Statistics*, 38: 489-497, 2015.
- 10) Kuhn, M. Building predictive models in R using the caret package. *Journal of statistical software*, 28(5):1-26, 2008.

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表
2. 学会発表

H. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

なし。

2. 実用新案登録

なし。

3. その他

なし。