

令和元年度厚生労働科学研究費補助金（政策科学総合研究事業）  
（総括）研究報告書

児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と  
AIを活用したリスク評価に向けた研究に関する研究

研究代表者 高岡 昂太 産業技術総合研究所

研究要旨：機械学習技術によりXGBoostを用いて保護判断履歴を高い精度で再現可能にしたことと同時に、誤予測の可能性を考慮したBaysian Deep Learningを用いて判断の補助を最大限補助できる可能性を考慮した。

A. 研究目的

本研究では児童の安全を継続的に確保するため、「虐待の再発(継続)」を考慮した対応判断や「一時保護の実施」に係る意思決定を支援する人工知能(以後AIと略す)開発と現場使用に関する探索的な研究を行う(研究1)。また、特に誤予測の可能性を減じるため、「誤予測」の可能性を考慮した技術設計と活用方法を確立する(研究2)。

B. 研究方法

それぞれ1自治体のデータベースを展開し、基本情報、リスクアセスメント、経過記録に関するデータを検討し、AIの研究開発を行った。

研究1の学習条件を次のように設定した。三重県の協力の下、学習データ80%(6924件)、検証データ20%(1731件)に無作為に分割。一時保護実施ラベルの不均衡(該当率14%)解消のためSMOTEサンプリングを適用した。eXtreme Gradient Boosting(XGBoost)モデルを利用する際に、5-fold 交差検証で学習roundを決定。損失関数は対数損失。学習率0.1を設定した。最大tree-depthは7(4~8を探索して最高精度)。その他デフォルト設定とした。予測対象は「一時保護の実施の有無」とし、リスク項目を含めた32項目の事例情報を採用した。なお、リスク項目は該当 = 1, 非該当と不明を0でラベルコーディングされ、通告経路等のカテゴリー情報はラベルエンコーディング(擬似連番)した。

研究2の学習条件を次のように設定した。学習データ80%(6924件)、検証データ20%(1731件)に無作為に分割とした。一時保護実施ラベルの不均衡(該当率14%)解消のためSMOTEサンプリングを適用した。採用する手法は、ベイズ多層ニューラルネットワークモデル(中間層3層 各36ノード)とした。予測対象は「一時保護の実施の有無」とし、リスク項目を含めた32項目の事例情報を採用した。リスク項目は該当 = 1, 非該当を0、不明を9でラベルコーディングし、通告経路等のカテゴリー情報はラベルエンコーディング(擬似連番)とした。

(倫理面への配慮)

産総研内の人間工学実験事前審査非該当、愛育研究所における倫理委員会の承認を得た。

C. 研究結果

研究1の結果は、図1・図2のように、閾値を0.5とする判別予測正解率(Accuracy) = 0.91 (91%) (クラス不均衡のため参考値)が得られた。

研究3の結果は、PR曲線下面積でXGBoostよりやや

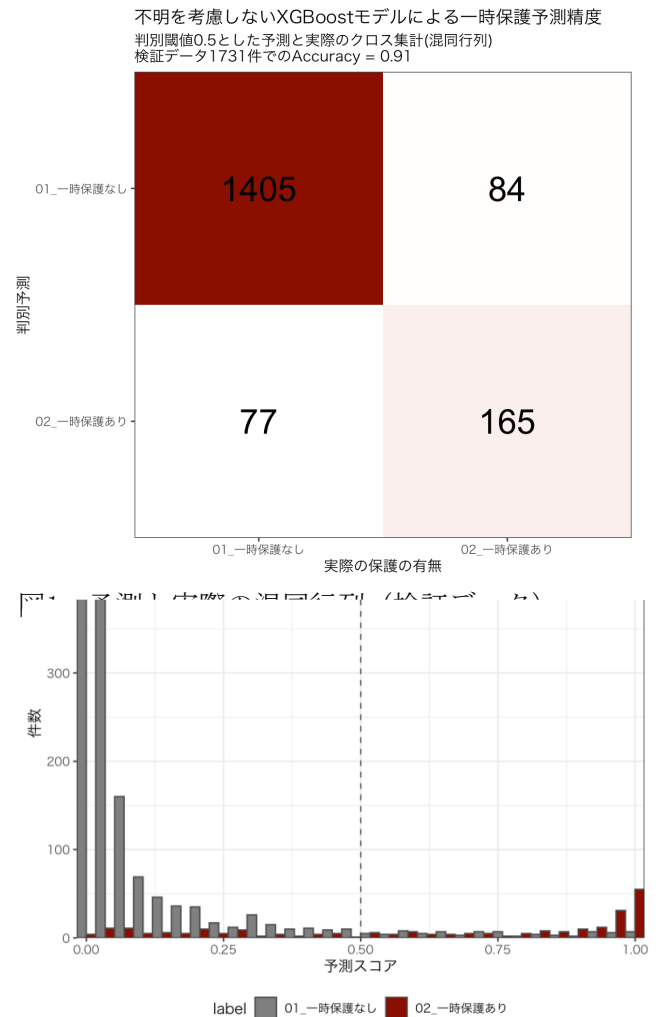


図2：実際の一時保護の有無別での予測スコア分布と設定閾値（検証データ）

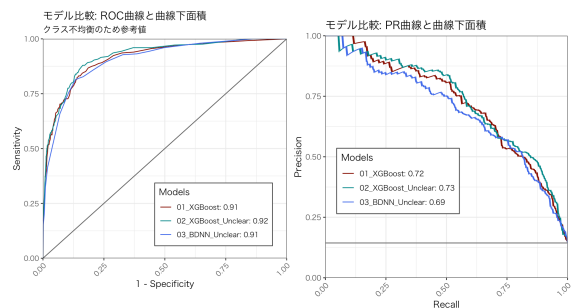


図3：各種性能指標

劣るものの、遜色ない性能を示した(図3)。また、幅が狭く(確信度が高く)予測が的中する場合と、幅が広く(確信度が低く)予測が誤っている場合を示すことができた。(図4)。

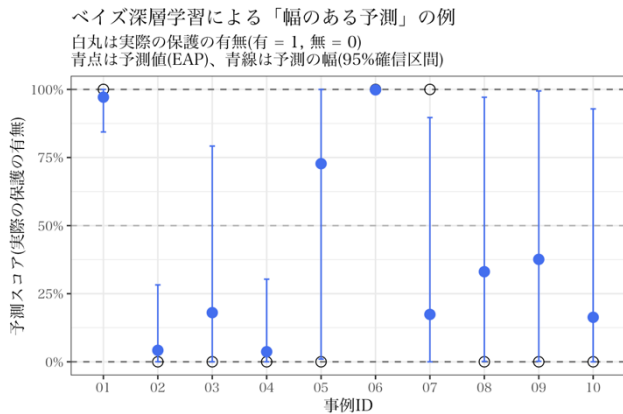


図4：事例毎の確信区間に関するイメージ

一方で、予測の幅が広いほど、誤予測率が高い傾向を事前に把握できた(図5)。

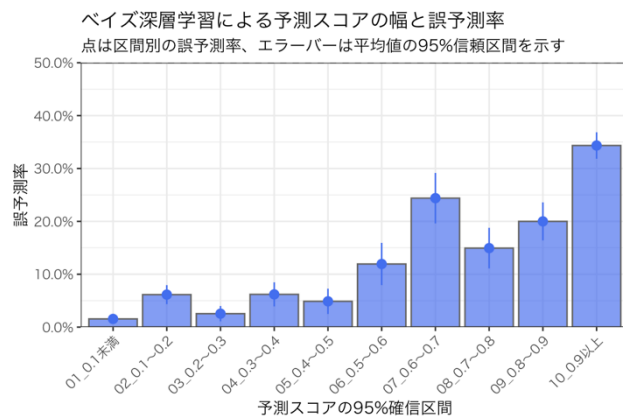


図5：誤予測率の把握

#### D. 考察

研究1については、XGBoostによりミスの少ない、一貫した一時保護判断が実現できたと考えられる。一方で、過去の一時保護判断が常に妥当であったとは限らないため、“要保護”に相当する情報を予測対象とした研究が必要(本事業範囲外)と考えられる。即ち、新規の虐待通告事例に対して、「過去ではその事例を一時保護していたか」を高い再現率で数的に評価可能となったと言える。

研究2については、図4のように、各新規事例に対する確信度付きの予測(予測区間)が実現された一方で、図5のように、確信区間の幅の大きさによって、どの程度の割合で当該予測結果が「誤りである

か」が事前に把握できるようになると考えられる。例えば、一時保護予測スコアが”0.9”と出力されても、予測の幅が0.9以上の場合(図5の10\_0.9以上に該当)、35%の割合で誤予測の可能性があると把握できる。一方で、確信区間の幅が0.1未満の場合、誤予測は2%未満とも言える。これにより、一時保護判断という重大な意思決定場面で、各事例に対する「機械学習による予測」がどの程度の確信度を持つものか、誤予測の可能性を事前に評価可能となったと言える。

こうした成果は、増加する虐待通告や対応職員の入れ替わりに伴って懸念される一時保護判断の一貫性の低下に対して、保護判断の水準を維持する客観的な指標が提供できる可能性が示唆された。

他にも、「一時保護の実施を予測する」という当該取り組みは“要保護児童の識別”に寄与すると考えられ、将来的にはコールセンター(区分対応システム)等での活用も期待されるだろう。

一方で、本研究の限界として、過去の一時保護判断が常に妥当であったとは限らない。そのため、本研究と同時並行に、“要保護”に相当する情報を予測対象とした研究が必要となる。次に、今回のAI研究開発は、データ提供元の特定の自治体に適用が限定される。個別自治体の資源(一時保護所の状況等)に即している点は適切だが、他の自治体に展開するためには、個別自治体でのデータ蓄積と解析が次年度以降の必要となる。最後に、誤予測の可能性を考慮したベイズ深層学習の精度(AUCPR)がやや低いことがあげられる。そのため、精度向上を検討すると、(解析的に大規模な)パラメータチューニングが必要となるだろう。

以上のように、データ提供元のように、初期初動のリスクデータが適切に収集されることで、重要なベースライン情報となり現場に必要なAI開発に役立てることが明らかになった。一方で、AIはあくまで一つの手段であるため、現場にデータ解析の結果を現場にフィードバックし、必要な業務改善や人材育成につなげることも同時に行っていく必要があるだろう。

#### E. 結論

一時保護判断については、ある程度精度が保てる結論となった。一方で、今後はAI以前に、良質なデータを集めるために、研修を含めた実務と平行した取り組みが不可欠である。

#### F. 健康危険情報

特になし

#### G. 研究発表

1. 論文発表：なし
2. 学会発表：なし

#### H. 知的財産権の出願・登録状況

1. 特許取得：なし
2. 実用新案登録：なし
3. その他：なし