

令和2年度厚生労働科学研究費補助金（政策科学総合研究事業）  
（分担）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：自治体の児童相談所がもつリスクアセスメントデータを元に、不明が多く選ばれた場合でも、ある程度精度を維持できるAI構築を行った。

分担研究者：北村光司  
産業技術総合研究所人工知能研究センター  
主任研究員

A. 研究目的

リスク項目への該当状況が「不明」の場合への対応。

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。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。解析アルゴリズムにeXtreme Gradient Boosting (XGBoost)モデルを利用し、5-fold 交差検証で学習roundを決定した。損失関数は対数損失とし、学習率0.1とした。最大tree-depthは7(4~8を探索して最高精度)とし、その他はデフォルト設定を採用した。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)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

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)を保持していた。“不明”のパターンを学習させても性能は維持されていることを確認した(図6, 図7参照)。

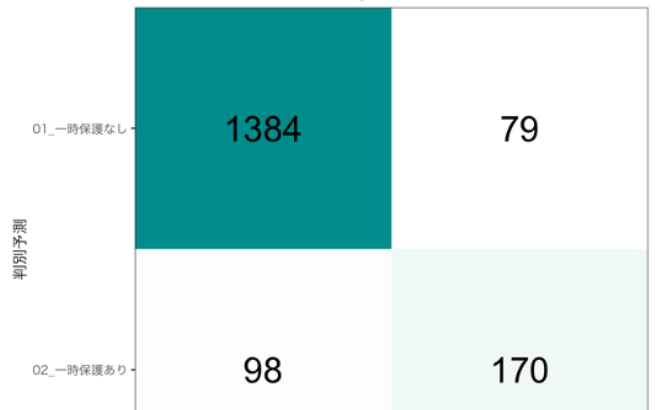
D. 考察

児童虐待対応現場において、初期初動体制では情報が不確かであったり、情報が集まらない場面が多くある。しかしながら、今回の調整を通して、初期初動で「不明」があったとしても、ある程度の精度を保てるAIを構築できたことは非常に大きい。さらに、今後は汎用可能性を含めて検討していくことが重要である。

E. 結論

初期初動において、リスクアセスメント項目に「不明」が回答されても、必要な一時保護を提案できる可能性を検討できた。今後は、汎用可能性があるか、さら

不明を考慮したXGBoostモデルによる一時保護予測精度  
判別閾値0.5とした予測と実際のクロス集計(混同行列)  
検証データ1731件でのAccuracy = 0.91



不明を考慮したXGBoostモデルによる予測スコア  
破線は閾値0.5を示す

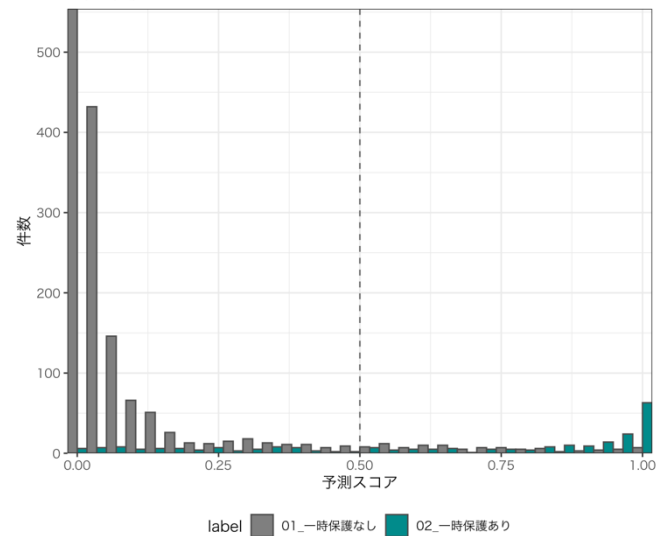


図7：実際の一時保護の有無別での予測スコア分布と設定閾値（検証データ）

に他の自治体でもデータ収集を含めた検討をしていくことが重要である。

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  
(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