

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

大規模データを利用した予測モデルの構築手法選択に関する研究
研究分担者 井上 永介
聖マリアンナ医科大学 医学部医学教育文化部門（医学情報学） 教授

研究要旨

医学的イベントの予測問題は古くから統計的手法で対処されてきたが、近年、人工知能を適用した研究成果の報告が多くなされている。本研究では、実例を通して、人工知能的手法にどの程度のメリットがあるのかを調査した。具体的には、実際の疫学研究データを用いた研究を調査した（Weng et al., PLOS One, 2019）。統計的手法（Cox 回帰モデル）・機械学習的手法（ランダムフォレスト）・人工知能的手法（深層学習）の死亡イベント予測性能を ROC 曲線の曲線下面積で評価したところ、それぞれ、0.75、0.78、0.79 であった。小児医療情報収集システムでも人工知能的手法の活用が求められるが、結果の確認とどの程度の利点があるかを数値化するために、統計的手法を並列して用いることが必要であろう。

A．研究目的

医学分野において、人工知能手法を適用した事例が増えてきている。FDA が眼科疾患の診断を行うデバイスを承認したことに始まり、人工知能による皮膚科疾患の分類性能は皮膚科医と同等というものや、頭部 CT スキャンの読み取り、乳がん検診画像を用いた診断補助など、画像を用いたものを中心として広く応用がなされている。画像を利用したものが多のは、深層学習の特性と医学分野で解決すべき課題が重なったことによるものと思われる。しかしながら、医学分野には、画像以外にも人工知能アルゴリズムを適用して解決すべき問題が数多くある。

本研究班で整備が進められている小児医療情報収集システムでは、相当な規模のデータベースが整備され、様々な分野での活用が期待されている。一つの適用例として、医学的イベントの予測というものがある。人工知能手法を利用すると、標準的な統計学的手法では達成しえなかった精度（ROC-AUC など）の予測が可能となり、実際、そういった報告が増えてきている。

予測に関する問題は、古くから統計学的モ

デルを利用したアプローチがとられてきた。統計モデルを利用するとある程度枠組みを固定（統計的にモデル化）することになるため、人工知能の学習に必要なデータ量よりは少なく済む。しかし、データを収集することの価値はすでに広く認識されており、大規模データベースは数多く存在する。このような環境では統計的モデリングにこだわる必要は無く、人工知能手法を積極的に適用すべきであろう。

本研究では、医学的イベントの予測において、予測手法の比較を行った文献を調査し、小児医療情報収集システムで実用的な手法選択について探ることを目的とする。

B．研究方法

2019年に発表された統計的手法と人工知能的手法を比較した文献（Weng et al., PLOS One, 2019）を調査し、予測性能の違いを実例から評価することとした。この結果をもとに、小児医療情報収集システムで人工知能的手法は適用可能であるか、適用可能であればどの程度の予測性能向上が見込めるかについて考察した。

(倫理面への配慮)

本研究には、倫理的考慮を必要とする内容は含まれていない。

C．研究結果

調査した文献は、2006年から2010年にUK Biobankに登録された40から70歳の502,628人を対象とした研究である。死亡をイベントとし、観察は2016年まで行われた。観察期間中に14,418イベントを観察した。比較する手法は人工知能的手法である深層学習(DL)、機械学習的手法のrandom survival forest(RF)、標準的統計手法のCox回帰モデルである。評価指標はROC曲線の曲線下面積(AUC)である。この指標は、1に近いほど分類能力が高いと解釈することができる。Cox回帰モデル、RF、DLのAUCは、それぞれ、0.751、0.783、0.790であった。これをもとに、著者らは、RFとDLは統計的フレームワークの手法(Cox)よりも予測能力が高い、と結論している。統計的手法に対するRFとDLのAUC増加量は0.04ほどである。加えて、RFとDLでは予測確率が正確(well calibrated)であった。

D．考察

死亡イベントを予測する問題でAUCが0.8程度という数値は高く、人工知能的手法でなければ達成できなかったものと考えられる。14,000件以上のイベント数を考えると、統計的手法であっても相当数の変数を扱うことができるため、AUCとして0.75という高い数値を達成している。だが、予測変数とイベントの間の関係を線形関係に定められてしまうため、この点でCoxはRF・DLに及ばなかったものと考えられる。

小児医療情報収集システムのサンプルサイズおよびイベント数はここで調べた研究と同等規模である。となると、本研究班でも予測問題を扱う場合は、人工知能的手法の適用を積極的に考慮すべきである。しかし、Wengらの研究のように、これまで長く用いられてきた統計的手法と比較することも大切である。こ

れにより、科学的に結果を比較・確認することができ、深層学習によるメリットがどの程度あったのかを数値化することができる。

E．結論

実際の疫学研究データを用いて、統計的手法に対する人工知能的手法の有用性が確認された。当然ではあるが、大規模データが利用可能である場合、人工知能的手法を利用することが望ましい。しかし、結果の妥当性確認およびメリットの数値化のため、従来手法である統計モデルによるアプローチも並行して実施するべきである。

F．健康危険情報

該当なし

G．研究発表

1. 論文発表

該当なし

2. 学会発表

該当なし

H．知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし