

東京大学医学部附属病院
株式会社島津製作所

人工知能により患者データから肝がんの存在を予測 -患者データからがんの存在を予測する AI の開発-

東京大学医学部附属病院と株式会社島津製作所は、「人工知能により患者データから肝がんの存在を予測」と題した論文を5月30日に英学術誌「*Scientific Reports*」にて発表しましたのでご報告いたします。

1. 発表者：

佐藤 雅哉 (東京大学医学部附属病院 検査部 助教)
建石 良介 (東京大学医学部附属病院 消化器内科 特任講師)
小池 和彦 (東京大学医学部附属病院 消化器内科 教授)
矢富 裕 (東京大学医学部附属病院 検査部 教授)
森本 健太郎 (島津製作所基盤技術研究所 AI ソリューションユニット 主任)
梶原 茂樹 (島津製作所基盤技術研究所 AI ソリューションユニット 主幹研究員)

2. 発表のポイント：

- ◆患者データに応じて予測能を最大化させる最適な機械学習アルゴリズム(注1)と学習パラメーター(注2)を自動抽出するフレームワーク(注3)を開発しました。
- ◆本研究結果により従来の腫瘍マーカー単一ではなく、臓器の炎症や患者背景などの情報も統合してがんの存在予測を行うことが可能になります。
- ◆日常の診療で収集される患者情報の有効利用や、患者情報を用いた診断支援への応用が期待されます。

3. 発表概要：

多要因が組み合わさり発症するさまざまな“がん”に対し、単一腫瘍マーカーでの存在予測には限界があり、患者背景や臓器の炎症などの情報も統合することが望ましいと考えられます。情報技術に大きな進展をもたらしたニューラルネットワーク(注4)を用いたディープラーニング(注5)の登場により近年注目される機械学習(注6)は、複数因子を組み合わせる際に関数の最適化を行い、予測能を最大化させるアルゴリズムを作成することを可能とします。

東京大学医学部附属病院 検査部の佐藤雅哉 助教、矢富裕 教授、同院 消化器内科の建石良介 特任講師、小池和彦 教授らおよび島津製作所基盤技術研究所 AI ソリューションユニット 梶原茂樹 主幹研究員らの研究グループは、ディープラーニングを含むさまざまな手法から、収集された患者データから得られる予測能を最大化する学習アルゴリズムと学習パラメーターを自動抽出するフレームワークを作成し、患者データを用いた肝がんの有無の予測精度を検討しました(図1)。

肝がんの有無の予測には、腫瘍マーカーの他、背景肝の線維化や炎症、肝炎ウイルスの有無、患者年齢などが重要であることが知られています。これらに関連する検査項目を含めた16項

目の患者データを、本フレームワークを用いて統合することで、従来の腫瘍マーカーと比較して診断率が飛躍的に向上しました。

ディープラーニングは多くの分野で革新的な成果をもたらした非常に強力な学習アルゴリズムですが、その複雑さのために大きな成果を生み出すには莫大な量のサンプルが必要になります。今回 1582 人の患者（肝がん患者 539 人、非肝がん患者 1043 人）を対象とした肝がん予測に対しての最適アルゴリズムはディープラーニングではない、従来の手法でした。

患者を対象とする医学研究においては、同意取得の必要性や倫理的な側面への配慮から、多数の（数万人の）患者サンプルを収集することは現実的に困難です。このような状況の中で、現存するデータに対して予測能を最大化するアルゴリズムを抽出することは非常に重要です。本フレームワークは肝がんに限らず、さまざまなデータに適用が可能であり他分野への応用も期待されます。

本研究成果は日本時間の 2019 年 5 月 30 日に学術誌 *Scientific Reports*（オンライン版）にて発表されました。

4. 発表内容：

【研究の背景】

多要因が組み合わさり発症するさまざまな“がん”に対し、単一腫瘍マーカーでの存在予測には限界があります。従来がんの有無の予測に使用される腫瘍マーカーは有効な手段ですが、日常の診療においては腫瘍マーカー以外にもたくさんの情報が収集されるため、なるべく多くの情報を統合して診断を行うことが望ましいと考えられます。情報技術に大きな進展をもたらしたニューラルネットワークを用いたディープラーニングの登場により近年注目される機械学習は、複数因子を組み合わせる際に関数の最適化を行い、予測能を最大化させるアルゴリズムを作成することを可能とします。人工知能を用いた画像認識においては、画像の画素それぞれの輝度情報を並べた数値の羅列のパターンをコンピューターが学習し、画像の分類を行っていますが、本手法は日常診療にて収集される患者データにも応用することが可能です。

本研究においては、さまざまな機械学習の中から入力したデータに応じて最も有効な予測を行うことができる学習アルゴリズムと学習パラメーターを自動抽出するフレームワークを作成し、日常の診療で得られる患者データを用いてどの程度正確に肝がんの有無が予測出来るかを検討しました。

【研究の内容】

1. 機械学習を用いたフレームワークの作成

研究グループは、線形回帰モデル（注 7）、ニューラルネットワーク、決定木（注 8）、サポートベクターマシン（注 9）、ディープラーニングといった様々な手法から、収集された患者データから得られる予測能を最大化する学習アルゴリズムと学習パラメーターを自動抽出するフレームワークを作成しました。本フレームワークに目的とする予測対象（本研究では肝がんの有無）と日常診療で得られる患者データセットを入力すると、本フレームワーク内で最も予測精度の高い学習アルゴリズムと学習パラメーターが抽出され、最適な学習モデルが自動的に作られます。

2. 肝がんの予測モデルの作成と精度の評価

1997 年 1 月から 2015 年 5 月までに東京大学医学部附属病院を受診した肝疾患患者の中から、肝がんの予測モデルに入力する年齢、性別、身長、体重、HBs 抗原と HCV 抗体（注 10）、ア

ルブミン、ビリルビン、AST、ALT、ALP、 γ GTP、血小板値（注11）、AFP、AFP-L3分画、PIVKA-II（注12）の16項目と”肝がんの有無”の情報が収集可能であった1582人（肝がん患者539人、非肝がん患者1043人）を用いて肝がんの予測モデルの作成と精度の評価を行いました。

対象患者1582人を、各アルゴリズムの訓練を行うための訓練データ（1266人）、最適な学習パラメーターを抽出するための検証データ（158人）、作成された学習モデルの精度を検証するための評価データ（158人）の3つにわけて検討を行いました。機械学習のフレームワークにより抽出された最適アルゴリズムは勾配ブースティング決定木（注13）で、 $\text{eta} = 0.08$ 、 $\text{gamma} = 0.02$ 、 $\text{max depth} = 1$ 、 $\text{minimal_child_weight} = 1.5$ 、 $\text{nround} = 300$ 、 $\text{subsample} = 0.5$ 、 $\text{colsample_bytree} = 0.9$ （注14）が最適パラメーター値として抽出されました。本アルゴリズムと最適パラメーターを用いた評価データにおける肝がんの正診率、AUROC値（注15）はそれぞれ87.3%、0.940と高い精度を示しました。本学習モデルは患者年齢や腫瘍マーカー、アルブミン値などといった因子に注目し、肝がんの有無の判定を行っていました（図2）。ディープラーニングは本データに対する最適なアルゴリズムではなく、正診率とAUROC値は83.5%、0.884でした。腫瘍マーカーを単独で用いた際の正診率とAUROC値はそれぞれAFPで70.7%、0.766、AFP-L3分画で74.9%、0.644、PIVKA-IIで71.1%、0.683でした。

【今後の展望】

現在、電子カルテのクラウド化などの試みも進んでおり、将来的にプラットフォームのクラウド化などにより症例集積が加速すれば本システムの更なる精度の向上が可能であると考えられます。患者を対象とする医学研究においては、同意取得の必要性や倫理的な側面への配慮から、多数の（数万人の）患者サンプルを収集することは現実的に困難です。このような状況の中で、現存するデータに対して予測能を最大化するアルゴリズムを抽出するというアプローチはとても重要と考えられます。本フレームワークは肝がんに限らず、さまざまなデータに適用が可能であり他分野への応用も期待されます。

5. 発表雑誌：

雑誌名：*Scientific Reports*（オンライン版：5月30日）

論文タイトル：Machine-learning Approach for the Development of a Novel Predictive Model for the Diagnosis of Hepatocellular Carcinoma

著者：佐藤雅哉*、森本健太郎、梶原茂樹、建石良介、椎名秀一朗、小池和彦、矢富裕（*責任著者）

DOI番号：10.1038/s41598-019-44022-8

6. 用語解説：

(注1) アルゴリズム：

コンピューターで計算を行うときの計算方法。

(注2) 学習パラメーター：

各計算手法の挙動を制御する設定値や制限などのルールのこと。

(注3) フレームワーク：

システムを構築するための基盤となる汎用的な機能を提供するプログラムの枠組みのこと。

(注4) ニューラルネットワーク：

脳の神経細胞ネットワークを模倣し、脳機能の特性をコンピューターで実現しようとしたもの。

(注5) ディープラーニング：

ニューラルネットワークの層を増やすことで、主に画像認識などの処理能力を画期的に向上させた人工知能の一つの形態。

(注6) 機械学習：

コンピューターがデータを反復的に学習し、そこに潜む特徴やルールを学ぶこと。

(注7) 線形回帰モデル：

求めたい結果 Y を実際に得られている事象 X から $Y=AX+B$ のような式で求める数理モデル。

(注8) 決定木：

求めたい特徴をよく反映するデータを順に見つけ、木の枝のように分類ルールを生成し、予測結果を分類する機械学習の手法の一つ。

(注9) サポートベクターマシン：

識別を行う境界と各カテゴリー間の距離にあたるマージンをなるべく大きくする「マージン最大化」という基準を用いて学習を行う機械学習の手法の一つ

(注10) HBs 抗原、HCV 抗体：

B型肝炎やC型肝炎ウィルスの有無に関わる採血項目。

(注11) アルブミン、ビリルビン、AST、ALT、ALP、 γ GTP、血小板値：

肝がんの炎症や線維化、肝臓の機能に関わる検査項目。アルブミンは栄養状態も反映しており、がんの病態が進行すると低下する。

(注12) AFP、AFP-L3 分画、PIVKA-II：

肝癌で上昇することが知られている（上昇しない肝癌もあります）腫瘍マーカー。

(注 13) 勾配ブースティング決定木:

小さな決定木を作成し、次の小さな決定木でそれを修正するように学習を進める決定木の手法の一つ。

(注 14) eta、gamma、max depth、minimal_child_weight、nround、subsample、colsample_bytree:

勾配ブースティング決定木に用いる学習パラメーターで、決定木の深さや学習の重みなどのルールを数値で示したもの。

(注 15) AUROC 値:

The area under the receiver operating characteristic curve の略で、判別モデルの性能を評価する指標の 1 つ。AUROC 値は 0 から 1 までの値をとり、値が 1 に近いほど判別能が高いことを示す。

7. 添付資料：

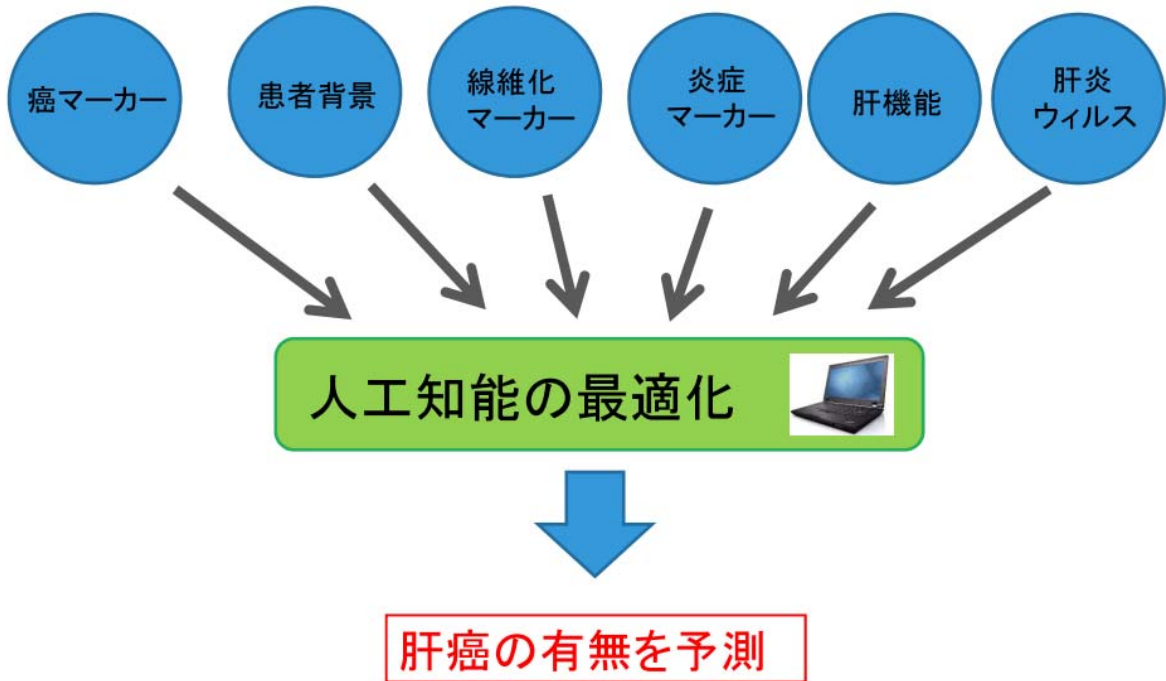


図 1: 入力された患者情報から、肝がんの予測能を最大化する最適な人工知能と学習パラメーターが適用され、未知の患者に対する肝がんの予測が行われます。

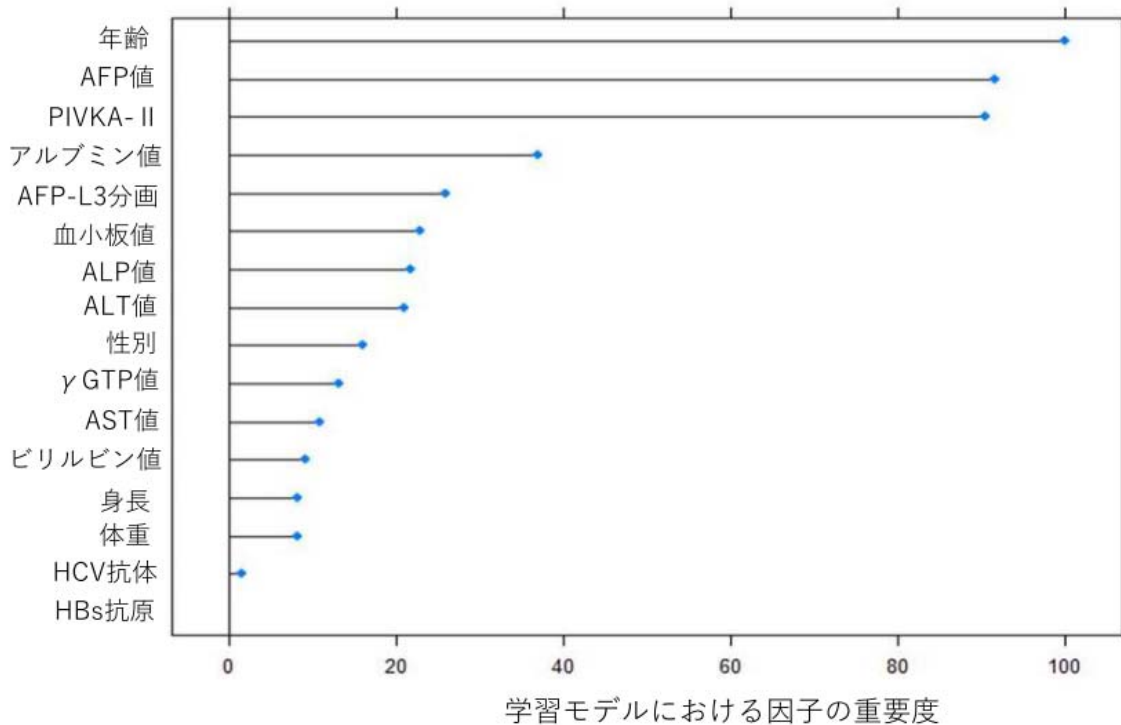


図 2: 最適化された学習モデルが肝がんの有無の予測に際して注目した因子の重要度を示したものです。年齢や腫瘍マーカー、アルブミン値などを重要視しており、実臨床における専門医の感覚に近いものとなっております。