

一般口演 | 医療データ解析

一般口演11 機械学習

2021年11月20日(土) 09:10 ~ 11:10 G会場 (2号館3階232+233)

[3-G-1-04] 統合診療データを用いた AIによるアウトカム予測と診療サマリ生成に向けた検討

*古川 大記¹、大山 慎太郎¹、佐藤 菊枝¹、狩野 竜示²、三沢 翔太郎²、鎌水 大和²、白鳥 義宗¹（1. 東海国立大学機構名古屋大学 医療健康データ統合研究教育拠点(C-Hit), 2. 富士フイルム株式会社）

*Taiki Furukawa¹, Shintaro Oyama¹, Kikue Sato¹, Ryuji Kano², Shotaro Misawa², Hirokazu Yarimizu², Yoshimune Shiratori¹（1. 東海国立大学機構名古屋大学 医療健康データ統合研究教育拠点(C-Hit), 2. 富士フイルム株式会社）

キーワード：integrative data analysis, machine learning, CITA

[背景と目的]

診療文書の作成にかかる医師の負担は大きく、特に週1回程度の診療サマリ作成業務は負担が大きい。また近年、院内診療データを用いた院内死亡予測や退院予測による急変予知や早期対応、治療効果予測による個別化医療の推進が期待されている。

このような文書作成や診療予測の支援を行うためには、院内診療データを統合して用いることが必要である。本研究では、統合診療支援プラットフォーム CITAで集約された院内データを用いて、院内死亡など複数の臨床的に重要なアウトカムを予測するアルゴリズムを作成し、臨床的に重要な要素を自動抽出して診療サマリを自動生成し臨床医に提示するシステムの構築を検討している。

そこで PoCとして、全診療科で診療する急性期疾患かつ日本人の死亡原因第5位である肺炎を対象に、臨床的なアウトカムを予測する AIアルゴリズムを構築する。

[方法]

2018年11月から2020年12月に名古屋大学医学部附属病院に入院した肺炎患者について CITAから統合診療情報を抽出した（患者数: 722例）。得られたデータについて、ランダム分割により症例毎のトレーニングとテストセットを作成した。その後、入院当日と入院初期評価時で各種予測を行う AIアルゴリズムを構築した。フリーテキスト形式のデータはベクトル化して解析を行った。予測アルゴリズムは AUCを用いて精度を評価した。

[結果]

予測アルゴリズムに用いる入力データ種別を組み合わせる事で、臨床的に有用な予測子を同定した。同定された予測子は臨床医学上も重要な因子であった。

[結語]

統合診療データを用いる事で、臨床的に重要なアウトカムを予測する機械学習モデルを構築し、臨床的に重要な予測子を同定した。得られた予測子は臨床医学上も重要な項目であり、診療サマリの自動生成アルゴリズムの構築可能性が示唆された。

統合診療データを用いた AI によるアウトカム予測と 診療サマリ生成に向けた検討

古川 大記^{*1}、三沢 翔太郎^{*2}、大山 慎太郎^{*1}、佐藤 菊枝^{*1}、狩野 竜示^{*2}、鎌水 大和^{*2}、白鳥 義宗^{*1}

^{*1} 東海国立大学機構名古屋大学 医療健康データ統合研究教育拠点 (C-HiT)、

^{*2} 富士フイルム株式会社

Outcome prediction using Integrated Clinical Database toward Automatic Generation of Clinical Summaries

Taiki Furukawa^{*1}, Shotaro Misawa^{*2}, Shintaro Oyama^{*1}, Kikue Sato^{*1}, Ryuji Kano^{*2}, Hirokazu Yarimizu^{*2}, Yoshimune Shiratori^{*1}

^{*1} Center for Healthcare Information Technology (C-HiT), Nagoya University, Nagoya, Japan, ^{*2} FUJIFILM Corporation

Outcome prediction using clinical data such as mortality prediction, length-of-stay prediction is applicable to acute change prediction, early treatment, and prediction of treatment effect. For the outcome prediction, integrating a large number of clinical data is indispensable. In this paper, we use a variety of clinical data aggregated in "CITA Clinical Finder", the Integrated medical support platform, to train mortality prediction and length-of-stay prediction models. For evaluation, we use patient data of pneumonia: the disease that is diagnosed in all services and the fifth leading cause of death in Japan. We conduct experiments on gradient boosting decision tree and logistic regression models. The result shows our model can predict more accurately than A-DROP, an indicator widely used to measure the seriousness of pneumonia. Also, the result shows that the model trained with data of all diseases can accurately predict the outcome of pneumonia cases, implying our model can be applied to outcome prediction of various kinds of diseases. We further discuss if the experimented models can extract clinically important information through the analysis of the contribution of each feature.

Keywords: Integrative data analysis, machine learning, CITA

1 はじめに

1.1 背景

近年、人工知能(AI)の発展に伴って、入院患者の院内診療データを用いた院内死亡予測や入院期間予測を行い、急変予知や治療効果予測による個別化医療の推進が期待されている。さらに、予後予測と入院期間予測に有用な臨床情報を抽出して使用する事で、単なる文書作成支援に留まらない、臨床上有用な診療サマリの作成支援ができる可能性がある。医師は、診療文書の作成に多大な労力を費やしており、こうしたシステムの実現は医師の負担を軽減すると期待される。

AIを用いた予後予測の研究はこれまでも取り組まれている。例えば、検査値やバイタルなどの数値データ、患者属性や診断名などのカテゴリーデータ、さらに診療記録のテキストを入力データとして用いた LSTM 深層学習モデルによる院内死亡率や入院期間、再入院の有無、診断名(ICD-9)の予測¹⁾や、入院時診療記録を入力とした BERT モデルによる予測が行われている²⁾。BERT モデルによる予測は、ICD コードの階層構造を統合することで、予後予測の性能が向上することを示している。これらの先行研究から、予測精度を向上させるためには特定のデータだけではなく、院内で記録される患者の多様なデータを統合して利用することが重要であることが示唆されている。

そこで我々は、富士フイルムメディカル IT ソリューションズ株式会社が開発する「統合診療支援プラットフォーム CITA Clinical Finder」(以下 CITA)で集約された院内データを用いて、院内死亡など複数の臨床的に重要なアウトカムを予測するアルゴリズムを作成し、さらに臨床的に重要な要素を自動抽出して診療サマリを自動生成し臨床医に提示するシステムの構築を検討している。本研究では Proof of Concept として、全診療科で診療する急性期疾患かつ日本人の死亡原因第5位である肺炎を対象に、AI による予後予測モデルと長期入院予測モデルを構築し、臨床現場における予測モデルの実

現可能性を評価する。次に、予測に有用な特徴量を分析することにより、予測モデルから臨床上有用な情報を抽出可能か分析する。

1.2 目的

本研究の目的は、多様な医療データを一元的に管理する統合診療支援プラットフォーム CITA に集約された院内データを網羅的に用いて、AI の一つである機械学習により、肺炎の 30 日以内死亡予測と長期入院予測を行う事である。作成した予測モデルは、肺炎の重症度モデルである A-DROP と予測精度を比較する。A-DROPは成人市中肺炎診療ガイドライン{日本呼吸器学会, 2017 年}でも採用されている肺炎の一般的な重症度モデルであり、元々は本邦の肺炎の死亡率に基づいて策定されたスコアである³⁾⁴⁾。

さらに、予測結果に寄与する特徴量を分析することでモデルが臨床上有用な特徴量を捉えられているかを確認する。

2 方法

2.1 本研究で実施する予測タスク

[死亡予測タスク] 入院初期における患者の死亡リスク予測は、治療場所、治療薬の選択、医療資源の配分などのトリアージにおいて重要な情報となる。肺炎患者においても 30 日以内の死亡予測は重要であり、一般的に A-DROP スコアが用いられている³⁾⁴⁾。そこで、各入院を1単位とし、入院時から 3 日以内の診療データを用いて、入院から 30 日以内に死亡する予測モデルを構築し、A-DROP スコアと予測精度を比較した。

[長期入院予測タスク] 肺炎は他の疾病に比べ入院期間が比較的長い³⁾ため、入院期間を予測する事で、医療リソースの適切な分配や治療方針の決定に寄与できる。そこで長期入院の予測評価で一般的に用いられる 15 日以上入院¹⁾

を長期入院と定義し、入院から 3 日以内の診療データを用いて、長期入院となる予測モデルを構築し、A-DROP スコアと予測精度を比較した。

[評価指標] 予後予測の評価に広く用いられているロジスティック回帰分析をベースとした AUROC (the Area Under the Receiver Operating characteristic Curve) を用いて精度評価を行った。検定は Delong's test⁵⁾ の両側検定を用い、 $P < 0.05$ を統計的に有意とした。

2.2 使用するデータの定義

[対象患者とデータ] 名古屋大学医学部附属病院の CITA プラットフォームを用いて、2018 年 11 月から 2020 年 12 月の間に入院診療を受けた 20 歳以上の患者を対象に、2017 年 12 月から 2020 年 12 月までの入院診療データを抽出し、後ろ向きに解析した。CITA プラットフォームに格納されている各部門の院内データをデータソースと定義した。データソースは電子カルテのテキスト記載情報(SOAP)、放射線検査報告書、入院診療計画書、臨床検査結果、食事情報、処置情報、入院情報を使用した。各データソースに含まれる情報は 2.3 に示す。

[学習データとテストデータ]

構築されたモデルの予測精度を、肺炎の重症度指標である A-DROP と比較するため、CITA プラットフォームに A-DROP が記載されている肺炎患者の入院データをテストデータとして用いた。

学習データは、テストデータ以外の 2 種類のデータを用いた;1 つは全入院データ("全疾患学習データ")、もう 1 つは肺炎で入院した患者データ("肺炎学習データ")である。

肺炎患者の定義は、入院診療計画書に記載されている病名に「肺炎」を含むものとした。A-DROP は、電子カルテのテキスト記載情報(SOAP)に記載されている値を用いて算出した。また、入院から 3 日以内に複数の異なる A-DROP が記載されている場合は、最後に記載された値を使用した。

2.3 予測モデルの特徴量設計

予後予測への寄与が想定された以下 7 種類のデータソースを利用した。具体的な特徴量設計の方法を以下に示す。

[臨床検査結果] 臨床検査結果項目ごとの検査値を特徴量として用いた。検査項目は、学習データにおいて 1 万回以上出現した 133 項目を用いた。

[食事情報] 食事の種類と食事量を用いた。データベース上にフリーテキスト形式で記録されている食事に関する情報から正規表現で食種のみを抽出し、頻出する 32 の食種を対象とした one-hot ベクトルを用いた。食事量は、食事を食べた割合を数値データとして用いた。

[処置情報] 処置の中で予後予測に寄与すると考えられた 30 種類の処置を特徴量として用いた。入院 3 日以内に実施された場合の値を 1 に、それ以外の場合を 0 とする one-hot ベクトルを用いた。

[SOAP] 医師診療録や看護記録などの SOAP テキストを用いた。テキストデータは構成単語を Word Embedding に変換し、文書種類ごとにベクトルの合計値を計算した。この Word Embedding は学習データに含まれる全ての SOAP データを FastText⁶⁾ を用いて学習した。

[放射線検査報告書] 主に CT 検査時に記載される放射線検査報告書の所見文のテキストを利用した。テキスト部分の処理は SOAP と同様の処理を行った。

[入院診療計画書] 入院診療計画書の中から、フリーテキストで記載されている検査計画、治療計画の項目を利用した。テキスト部分の処理は SOAP と同様の処理を行った。

[入院情報] 入院に関するメタ情報を特徴量として用いた。具体的には入院時年齢は数値として用い、入院時の診療科、病名、入院種別(予定入院か緊急入院)、性別は one-hot ベクトルに変換した。

2.4 モデルによる予測および分析

A-DROP スコア、ロジスティック回帰⁷⁾、勾配ブースティング決定木⁸⁾を用いて予測を行った。

[A-DROP スコア] 年齢、BUN、SpO₂、意識障害、血圧の各項目の陽性数により算出される合計スコアを用いた。取りうるスコアは 0~5 点であるが、0.2 倍して 0~1 に変換した。

[ロジスティック回帰 (LR)] 一般化線形モデルの一種で、特徴量から確率値を推定する。パラメータ推定ではニュートン法を用いて最適化を行った。実装には scikit-learn を用いた。

[勾配ブースティング決定木 (GBDT)] 機械学習のうち、決定木のアンサンブル手法の一種である。誤差に応じて予測値を更新するように決定木を増やして学習を進める。実装には lightGBM⁹⁾を用いた。

LR および GBDT は 2 種類の学習データ(全疾患学習データ、肺炎学習データ)を用いてモデルの学習を行った。加えて、GBDT については、学習データの不足に対処しつつ肺炎に特化した分類を行うため、肺炎学習データを除いた全疾患学習データを用いて事前学習し、続いて肺炎学習データで追加学習を行ったモデルを作成した。これらのモデルに対して、テストデータで精度を評価した。

作成されたモデルを用いて、学習データ毎、予測タスク毎に、GBDT の分岐条件ごとの特徴量寄与度の合計値を基準とした特徴量の寄与度を算出し、データソース毎と特徴量毎に集計した。本分析では特徴量間の相関は考慮していない。また、Word Embedding の解釈は困難であるため、分析対象から除外した。

3 結果

3.1 患者背景

対象期間内に入院した 20 歳以上の患者の入院データ 55,973 件中、入院から 3 日以内に A-DROP スコアが算出可能であった肺炎患者の入院データ 226 件をテストデータとして用いた。テストデータには学習で用いた患者の入院データは含まなかった。

対象期間内における肺炎入院患者の死亡率が低く、死亡予測モデルの学習が非効率的であるため、学習データは死亡:生存が 1:10 となるようにランダムサンプリングを行った。各タスクにおける学習データとテストデータの総数と死亡数・長期入院数を表 1 に示した。また、各データソースの欠損率を表 2 に示した。

テストデータにおける A-DROP の分布、平均年齢、男女比率を表 3 に示した。平均入院期間は A-DROP の値が高くなるほど長くなっていた。

3.2 死亡予測と長期入院予測モデル

学習データとして全疾患学習データを用いた場合と肺炎学習データを用いた場合の、死亡予測モデルと長期入院予

測モデルの精度を、A-DROPスコアと比較した。

30 日以内死亡予測 AUROC は、学習データによらず、GBDT は A-DROP より有意に予測精度が高かった(図 1、2、表 4)。また、LR に関しても A-DROP よりは予測精度が高いが、GBDT の方がより高い予測精度であった。

学習データごとの性能を比較すると、全疾患学習データを学習データとした方が有意に高い予測性能を示した。

長期入院予測でも同様に、学習データによらず、GBDT は A-DROP より有意に予測精度が高かった(図 3、4、表 5)。

表 1: 各タスクの学習データとテストデータの総数と死亡数／長期入院数

データ種類	タスク	総数	死亡数	長期入院数
全疾患学習データ	30 日以内死亡	4,620	420	-
	長期入院	55,211	-	12,451
肺炎学習データ	30 日以内死亡	440	40	-
	長期入院	655	-	276
テストデータ	-	226	11	86

表 2: 各学習データの入力情報の欠損率

データソース名	全疾患学習データ 欠損率 (死亡/長期入院)	肺炎学習データ 欠損率 (死亡/長期入院)
SOAP	0.01 / 0.01	<0.01 / <0.01
放射線検査報告書	0.99 / 0.99	0.97 / 0.98
入院診療計画書	0.57 / 0.56	0.61 / 0.63
検査	0.20 / 0.21	0.04 / 0.05
処置	0.81 / 0.83	0.57 / 0.59
食事種類	0.13 / 0.11	0.19 / 0.19
食事量	0.13 / 0.11	0.20 / 0.19
入院情報	<0.01 / <0.01	<0.01 / <0.01

表 3: テストデータの統計値

A-DROP	数	30 日以内死亡		長期入院		平均入院日数	平均年齢	男女比
		数	割合	数	割合			
0	28	0	-	8	0.29	13.1	62.2	18:10
1	80	3	0.04	20	0.25	14.1	69.9	56:24
2	67	0	-	27	0.40	17.6	75.3	47:20
3	43	7	0.16	27	0.63	21.2	80.3	33:10
4	8	1	0.12	4	0.50	24.4	79.6	4:4
5	0	0	-	0	-	-	-	-
合計	226	11	0.05	86	0.38	16.8	72.9	158:68

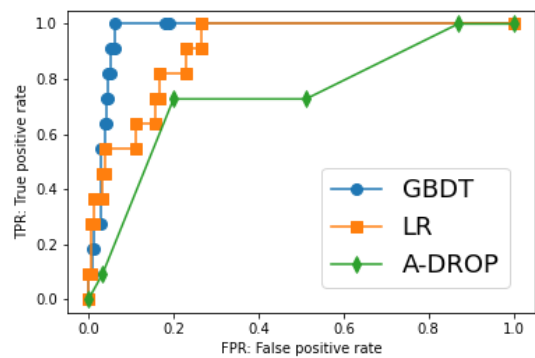


図 1: 全疾患学習データを用いた死亡予測の ROC 曲線

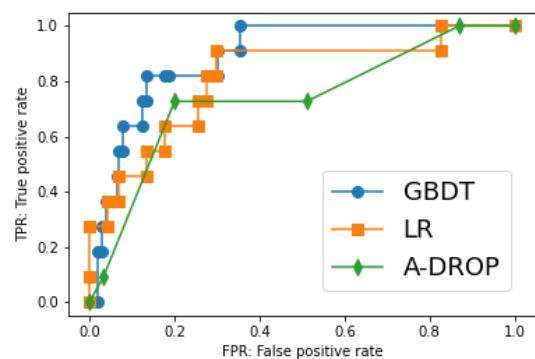


図 2: 肺炎学習データを用いた死亡予測の ROC 曲線

表 4: 学習データとモデル毎のテストデータに対する死亡予測精度

学習データ	モデル	AUROC	P 値*
肺炎	LR	81.10	0.2670
	GBDT	88.79	0.0238
全疾患	LR	90.70	0.0065
	GBDT	96.87	0.0060
	A-DROP	73.62	-

*A-DROP スコアに対する検定。

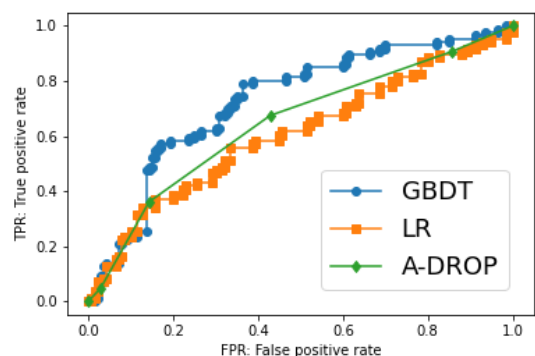


図 3: 全疾患学習データを用いた長期入院予測の ROC 曲線

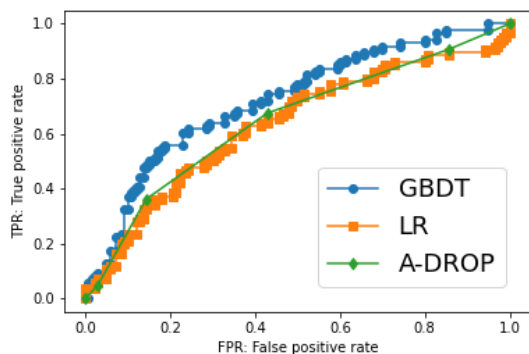


図 4: 肺炎学習データを用いた
長期入院予測の ROC 曲線

表 5: 学習データとモデル毎の
テストデータに対する長期入院予測精度

学習データ	モデル	AUROC	P 値*
肺炎	LR	63.29	0.7596
	GBDT	72.21	0.0438
全疾患	LR	60.81	0.3653
	GBDT	72.95	0.0245
	A-DROP	64.68	-

*A-DROP スコアに対する検定。

3.3 寄与度が大きい特徴量の分析

[寄与度の大きいデータソース] GBDT の分岐条件ごとの特徴量寄与度の合計値を基準とした特徴量の寄与度を、データソースと学習モデルごとに算出した結果を図 5 に示した。どの設定においても SOAP、検査結果、食事情報の寄与度が大きかった。一方、放射線検査報告書や処置情報は寄与度が小さかった。

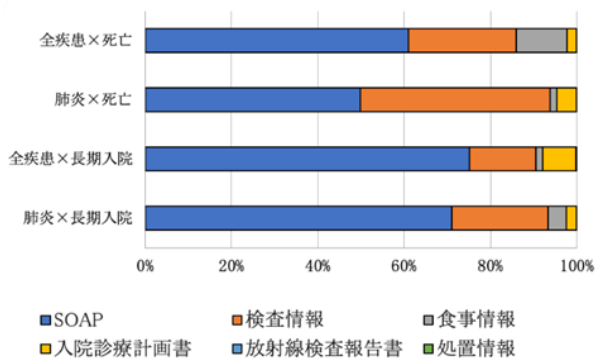


図 5: データソースごとの寄与度

[寄与度の大きい特徴量] 学習データと予測タスク毎に、寄与度が大きい特徴量を算出した(表 6-9)。どのタスクでも、検査情報の寄与度が大きかったが、特に死亡予測では、学習データに関わらず、尿素窒素、入院時年齢、アルブミンの寄与度が大きかった。長期入院予測では、全学習データを用いると入院科ラベルの寄与度が大きかったが、肺炎患者学習データを用いると食事量と血液検査データの寄与度が大きかった。

表 6: 全疾患の学習データを用いた
死亡予測モデルの寄与度上位 10 件

特徴量名	データソース	寄与度
副食量	食事情報	1636.2
尿素窒素	検査情報	740.3
乳酸脱水素酵素(LD)	検査情報	519.5
C 反応性蛋白(CRP)	検査情報	331.5
補正カルシウム	検査情報	261.7
アルブミン	検査情報	242.7
入院時年齢	入院情報	203.0
好中球%(分析機)	検査情報	143.2
呼吸器内科	入院情報	113.8
アルカリフォスファターゼ (ALP)	検査情報	102.6

表 7: 肺炎疾患の学習データを用いた
死亡予測モデルの寄与度上位 10 件

特徴量名	データソース	寄与度
アルブミン	検査情報	129.2
好中球%(分析機)	検査情報	103.6
尿素窒素	検査情報	98.0
カリウム	検査情報	37.5
リンパ球	検査情報	36.4
単球	検査情報	35.7
クロール	検査情報	31.9
入院時年齢	入院情報	28.1
主食量	食事情報	23.4
好中球数(分析機)	検査情報	21.9

表 8: 全疾患の学習データを用いた
長期入院予測モデルの寄与度上位 10 件

特徴量名	データソース	寄与度
整形外科	入院情報	5212.5
輸血前感染症用保存日	検査情報	3915.1
血液内科	入院情報	3461.8
心臓外科	入院情報	2535.3
消化器外科2	入院情報	2094.2
呼吸器外科	入院情報	1856.9
アルブミン	検査情報	1594.0
婦人科	入院情報	1198.2
C 反応性蛋白(CRP)	検査情報	1081.0
産科	入院情報	1009.6

表 9: 肺炎疾患の学習データを用いた
長期入院予測モデルの寄与度上位 10 件

特徴量名	データソース	寄与度
主食量	食事情報	119.1
副食量	食事情報	83.9
赤血球分布幅	検査情報	72.7
補正カルシウム	検査情報	44.4
プロカルシトニン(PCT)	検査情報	42.9
ナトリウム	検査情報	42.1
乳酸脱水素酵素(LD)	検査情報	41.7
平均赤血球容積	検査情報	39.6
リンパ球	検査情報	39.0
血漿中総二酸化炭素濃度	検査情報	34.4

3.4 学習データの違いによる性能比較

事前学習として全疾患学習データの中で肺炎以外のデータを用いて GBDT の決定木を学習し、続いて肺炎学習データで収束するまで GBDT の決定木を追加して追加学習を行ったところ(表 10)、死亡予測と長期入院予測において、肺炎学習モデルや全疾患学習モデルよりも高い予測精度を示した。

表 10: 学習データを変化させた際の AUROC

	死亡予測	長期入院予測
肺炎	88.79	72.21
全疾患	96.87	72.95
肺炎入院以外 →肺炎入院	97.76	73.51

4 考察

GBDT による死亡予測では、全疾患学習データを用いた場合でも肺炎に対して精度が高く予測できたことから、今回の学習モデルは汎用的なモデルを学習している可能性が示唆され、COVID-19 や他の疾患の患者に対する予測タスクに関しても高い予測精度が期待される。

GBDT による長期入院予測では、学習データの違いによる予測精度の有意差は認めなかった。これは、死亡症例が少ないため、死亡予測において学習データを肺炎学習データに限定すると学習データが不足するのに対し、長期入院の件数は多いため、学習データを肺炎学習データに限定した場合でも十分なサンプル数が得られたためと考えられた。また、死亡予測と比較して長期入院予測の予測精度は高くなかった。これは社会的な入院など、症例毎の医療以外の事情によって左右される事象であることも原因と考えられた。今後さらなる検討が必要である。

予測モデルの寄与度分析では、他の特徴量に比較して、放射線検査報告書や処置情報の寄与度が小さかった。これは比較的欠損率が高いため、構築したモデルがこれらの特徴量をあまり用いずに学習したのと考えられた。また、肺炎学習データでは全疾患学習データよりも検査データの寄与度が大きかった。これは特定の疾患で予測に寄与する検査結果があり、GBDT が重要な因子として絞り込みやすくなったことが一因と考えられた。さらに予測タスクを比較すると、死亡予測の方が検査結果の寄与度が大きかったが、死亡リスクの

高い患者は、特定の検査結果が顕著に悪化する傾向が強いと考えられた。

また、個別の特徴量の寄与度に関しては、どのタスクでも検査情報の寄与度が大きかったが、特に死亡予測に関しては、尿素窒素、入院時年齢、アルブミンといった従来から報告されている肺炎の予後と関連する因子の寄与度が大きかった。長期入院予測では、栄養状態を表すアルブミンや食事量の寄与度が大きかった。特に重症患者では異化亢進による栄養状態悪化により全身状態の改善遅延が起こるため、本結果は臨床上も妥当な結果と考えられた。さらに、肺炎学習データモデルと全疾患学習データモデルを比較すると、全疾患学習データモデルの方が診療科など幅広いデータに対応させるための特徴量の寄与度が大きかったが、肺炎学習データモデルの方がアルブミンや CRP、PCT の寄与度が大きく、急性感染症である肺炎患者に特化した特徴量の寄与度が大きかった。

GBDT による予測性能は全疾患のデータを利用したモデルの方が良い結果であった。これは、肺炎患者に限定した学習モデルは肺炎に特化しているものの学習データの数が不足するためであると考えられた。さらに、肺炎以外の全疾患学習データを用いて事前学習し、肺炎学習データで学習したところ、死亡予測と長期入院予測では最も高い予測精度を示した。このことから、肺炎患者を予測するためには、肺炎患者のデータを用いた方が良い可能性が示唆された。今後、肺炎患者のデータをより多く収集することで、さらなる精度改善が見込まれる。また、他疾患に適応する場合も同様のモデル構築手法が有用である可能性がある。

5 結語

今回我々は、統合診療支援プラットフォームである CITA を用いて、入院日から 3 日以内の情報から肺炎入院患者の 30 日以内死亡予測と長期入院予測を行う機械学習モデルを構築した。30 日以内死亡予測と長期入院予測の両方において、機械学習モデルは既存の重症度スコアである A-DROP を上回る予測精度を示した(AUROC スコア 97.76 vs 73.32%, 73.51 vs 64.68%)。構築した精度の高い予後予測モデルを用いることで、急変予知や早期対応に応用が可能であり、治療や処置を施した後の予測結果の変化から治療効果予測を行うことも可能になった。また、本研究で学習したモデルは他疾患への汎用性を有している可能性が示唆され、例えば COVID-19 患者に対する予測なども期待できる。

さらに、予測に有用な特徴量を分析したところ、SOAP や、処置、食事量など様々な種類のデータの統合が予測に有用であり、多様なデータを統合的に用いる事で臨床上も有用な予測が可能になると思われた。予測に重要な特徴量は臨床上も重要な因子であった。

今回抽出された予後予測と長期入院予測に有用な特徴量を用いて、診療サマリ作成が可能かさらなる検討が必要である。

参考文献

- 1) Betty van Aken, Jens-Michalis Papaioannou, Manuel Mayrdorfer et al. "Clinical Outcome Prediction from Admission Notes using Self-Supervised Knowledge Integration", Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, page 881—893. 2021.
- 2) Rajkomar, Alvin, Oren, Eyal, Chen, Kai et al, "Scalable and

- Accurate Deep Learning with Electronic Health Records”, npj Digital Med. Volume 1, 18. 2018.
- 3) 健山 正男, 新里 敬, 原永 修作, 他. A-DROP に基礎疾患と呼吸数を追加したシステムの 30 日死亡予測の検討. 日本呼吸学会誌 2011; 49(5): 343-348.
 - 4) 樋口多恵子, 太田 求磨, 田邊 嘉也, 他. 高齢者市中肺炎の重症度分類と予後予測. 老年医学会雑誌 2007; 44(4): 483-489.
 - 5) Elizabeth R. DeLong, David M. DeLong and Daniel L. Clark-Pearson. “Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach”, BIOMETRICS. Volume 44, page 837--845, 1988.
 - 6) Piotr. Bojanowski, Edouard Grave, Armand Joulin et al, “Enriching Word Vectors with Subword Information”, Transactions of the Association for Computational Linguistics. Volume 5, page 135-146. 2017.
 - 7) Raymond E. Wright, “Logistic regression” 1995.
 - 8) Ke Guolin, Qi Meng, Thomas Finley et al, “Lightgbm: A Highly Efficient Gradient Boosting Decision Tree”, Advances in Neural Information Processing Systems. Volume 30, page 3146-3154. 2017.