

一般口演

## 一般口演22

## 標準化・EHR・PHR

2018年11月25日(日) 13:40 ~ 15:40 B会場 (4F 409+410)

## [4-B-3-8] 臨床研究情報処理フレームワークにおける診療データの粒度調整

○松尾 亮輔, 小川 泰右, 荒木 賢二, 山崎 友義, 串間 宗夫 (宮崎大学医学部附属病院医療情報部)

【目的】本研究は、臨床研究における情報処理フレームワークを提案する。フレームワークは、データ前処理（診療データの粒度調整）と分析（群間比較と予測・分類）で構成し、これらの2フェーズの試行を繰り返すことで前処理の良し悪しをノウハウとして蓄積する。本発表では、肺血栓塞栓症予防管理料の算定・非算定の予測モデル構築において、診療データの粒度調整の有効性を検証する。【対象と方法】予測モデルの目的変数は肺血栓塞栓症管理料の算定有無の2値とし、説明変数はKコード、検査・観察項目とする。宮崎大学医学部附属病院の2015・2016年の2年分の様式1、検査・観察項目の診療データを用いる。患者数は計21,529名である。粒度調整の対象は、手術術式のKコードおよび数値データの検査・観察項目とする。粒度調整は、Kコードには重心法による階層クラスタリングを適用し、検査・観察項目には3分割の等間隔区間による離散化を行う。予測モデルの良し悪しの検討には5つの評価指標を用いる。診療データの粒度調整の有効性は、粒度調整あり・なしの場合において、5つの分類器による評価指標の平均値を用いて実験的に比較する。分類器のパラメータはデフォルトとし、訓練・テストデータを7割・3割とする。【結果】粒度調整を実施すると、粒度調整しない場合と比較して、AUCが0.52%増加、寄与率がほとんど変化なく、Precisionが0.9%減少、Recallが1.68%増加、F-measureが0.87%増加した。【結語】本研究は、臨床研究における情報処理フレームワークを提案し、フレームワーク内の診療データの粒度調整について、その有効性を検証した。実験結果から、粒度調整をするとPrecisionの適合率は下がったが、それ以外の評価指標では向上がみられた。PrecisionとRecallの総合的指標のF-measureが向上し、AUCの増加から算定患者かどうかを判別する精度が向上したことから、粒度調整の有効性が本実験結果から示唆された。

# 臨床研究の情報処理フレームワークにおける診療データの粒度調整

松尾亮輔<sup>\*1</sup>、小川泰右<sup>\*1</sup>、荒木賢二<sup>\*1</sup>、山崎友義<sup>\*1</sup>、串間宗夫<sup>\*1</sup>

<sup>\*1</sup> 宮崎大学医学部附属病院医療情報部

## Granularity Adjustment of Clinical Data for a framework of Information Processing on Clinical Researches

Ryosuke Matsuo <sup>\*1</sup>, Taisuke Ogawa <sup>\*1</sup>, Kenji Araki <sup>\*1</sup>, Tomoyoshi Yamazaki <sup>\*1</sup>, Muneo Kushima <sup>\*1</sup>

<sup>\*1</sup> Department of Medical Informatics, Faculty of Medicine, University of Miyazaki Hospital

We propose a framework of information processing on clinical researches. The proposed framework composes of two phases: data preprocessing of granularity adjustment of clinical data and analyses such as controlled trial, prediction and classification. The framework cyclically accumulates the feature importance of processed clinical data derived from the analysis phase as the evaluation of the data preprocessing. In this paper, we verify the effectiveness of granularity adjustment of clinical data such as the codes of surgeries and the data of laboratory tests and observations on a case study of the prediction of three management fees. Hierarchical clustering is then exploited for the surgeries' codes and the normal distribution is exploited for the data of laboratory tests and observations. Regarding the experimental evaluation of the granularity adjustment, we execute 5-fold cross validation with a paired t-test and four evaluation measures: AUC, Precision, Recall and F-measure. We employ five classifiers: Logistic Regression, Naive Bayes, Neural Network, Random Forest and SVM(linear). The experimental results using the data of 21,529 patients in University of Miyazaki Hospital showed that granularity adjustment generally brought about better performance of the prediction of three management fees in terms of Recall and F-measure.

**Keywords:** Granularity adjustment, clinical data, hierarchical clustering, discretization, prediction.

### 1. 緒論

医療分野の研究開発に資するための匿名加工医療情報に関する法律(次世代医療基盤法)が、2018年5月11日に日本で施行され、膨大な医療情報の利活用による健康・医療に関する先端的研究開発及び新産業創出の促進が期待されている<sup>1)</sup>。

医療情報は、数値・文字列・コードなど様々な形で表現されており、臨床研究や多変量解析などの後向き分析の目的に適した形に変換することが、分析の効率化に重要である。例えば、検査・観察項目データを分析する際は、数値データを離散化することで、教師あり学習の分類問題のような離散値での分析が必要な分析目的に対して効果的である。また、例えば、手術術式の K コードは部位ごとに構造化され、広く普及しているが、粒度が細かすぎる恐れがあることや、分析目的によっては部位以外の視点で構造化されていることが有効であることが考えられる。

機械学習を用いた Current Procedure Terminology(CPT)の粒度調整により、手術後の有害事象の予測精度が向上したという報告がある<sup>2)</sup>。また、ゲノム・プロテオームに関する24の生物医学データセットを用いて、機械学習による連続値の離散化により、分類精度が向上したという報告がある<sup>3)</sup>。血液ガス分析の検査値データセットと心拍出量の生理データセットを用いた分類精度が離散化により向上したという報告もある<sup>4)</sup>。

このように、診療データの粒度を調整する処理が予測や分類に重要であるが、分析に適切な粒度だけでなく、うまくいかなかった失敗結果も試行錯誤のプロセスとして蓄積できれば、前処理のノウハウが医療情報の二次利用に活用できる。また、前処理から分析までの一連のプロセスを方法論化できれば、過去の医療情報リアルワールドデータを用いた後向き分析の効率化が期待される。

### 2. 目的

本研究は、臨床研究における情報処理フレームワークを提案する。提案するフレームワークは、データ前処理(診療データの粒度調整)と分析(群間比較と予測・分類)の2つのフェーズで構成される。そして、2フェーズのサイクルにより、分析を実行することで得られる前処理後の診療データの特徴重要度を前処理の良し悪しの結果として蓄積できるものである。本研究では、肺血栓塞栓症予防管理料、輸血管理料1、麻酔管理料1(閉鎖循環式全身麻酔)の算定患者予測を事例として、診療データの粒度調整の有効性を検証する。

### 3. 方法

#### 3.1 フレームワークの概要

フレームワークの概要図は図1である。以降で2フェーズの詳細と2フェーズのサイクルによる前処理のノウハウの蓄積方法について述べる。

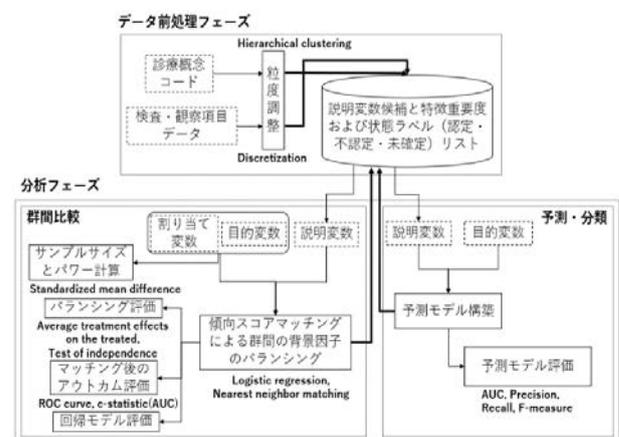


図1 フレームワークの概要

データの前処理フェーズでは、検査や観察項目における連続値データと手術術式や傷病などの診療概念コードの粒度を調整する機能を設計する。手術術式のKコードや傷病のICDコードのように既に階層構造で体系化されている場合は、階層クラスタリングを用いて、既存の体系をベースに、設定する分類軸により診療概念コードの体系を再構成する。採血データや、体温・血圧などの検査・観察項目の連続値データについては最も細かい粒度の連続値から離散化を行う。その際、粒度の設定や粒度調整の手法の選択では、後述の分析フェーズから得られるフィードバックをもとに、その良し悪しを識別する。

分析フェーズでは、後向き分析として例えば臨床研究としての群間比較や、多変量解析としての予測・分類の機能を設計する。群間比較の手順は、まず、サンプルサイズとパワー計算を行う。その後、ロジスティック回帰分析により傾向スコアを計算し、傾向スコアマッチングによる群間の背景因子のバランスを行う。最後に、回帰モデルの評価(C 統計量(AUC))、バランス評価(各説明変数の群間の標準化平均差)、アウトカム評価(処置群の平均処置効果と独立性検定)を行う。予測・分類では、複数の分類器の手法を用いて予測モデルを構築する。モデルの評価には AUC、Precision、Recall、F-measure の指標を用いる。複数の分類器により評価指標を求め、実験的に予測モデルの評価を行う。

そして、2つのフェーズのサイクルにより、前処理のノウハウを蓄積できる機能を設計する。ここでは、予測や分類といった分析結果から得られる説明変数の特徴重要度を活用することで、粒度調整がされた変数について、分析において重要とされる変数の粒度の調整の仕方を成功例として記録し、分析において重要とされない変数の粒度の調整の仕方を失敗例として記録していく。その際、粒度調整が成功した説明変数には状態ラベルとして「認定」を付与し、粒度調整に失敗した説明変数には状態ラベルとして「不認定」を付与する。そのほかの説明変数候補には「未確定」の状態ラベルを付与する。補足情報として、粒度調整の方法と分析の実験方法を記録していく。この機能により、前処理のノウハウを蓄積するだけでなく、人工知能技術を用いた説明変数候補の発見や、既存知識の検証が可能になると考える。

### 3.2 粒度調整

本稿の事例研究における粒度調整の対象は、手術術式のKコードおよび数値データの検査・観察項目である。粒度調整は、Kコードにはユークリッド距離を用いた重心法による階層クラスタリングを適用する。粒度調整のためのデータは様式1を用い、主傷病をKコードの階層クラスタリングの要素とする。その際、主傷病に該当するICD10コードの大文字のアルファベットのみを用いる。そして、部位ごとのKコードのまとまりに対して、逐次、階層クラスタリングを実行する。階層クラスタリングにはSciPy<sup>5)</sup>を用いる。表1は、階層クラスタリング結果の例である。

表1 階層クラスタリング結果の例

クラスターID	Kコード
(心、心膜、肺動静脈、冠血管等)1	K599-2, K552-21, K0541, K557, K0543, K598, K0542, K563, K0584, K570-3, K0562, K0583, K0561
(心、心膜、肺動静脈、冠血管等)2	K599-3

血管等)2	
(心、心膜、肺動静脈、冠血管等)3	K552-22
(心、心膜、肺動静脈、冠血管等)4	K597-2
(動脈)1	K610-3, K616-4, K613, K607-2
(動脈)2	K609-2
(動脈)3	K616

検査・観察項目は正常値・異常値の2分割の離散化を行う。入院患者を対象に正常値と異常値を判別するため、正規分布の活用により基準値を推定し、推定基準値内の値を正常値、推定基準値外の値を異常値とする。検査・観察項目どちらにおいても、文字列のデータが1箇所も含まれず、かつ、データの含有率が対象の全患者の5%以上の項目を離散化の対象とする。

検査データの離散化の対象は156項目であり、表2はデータの含有率の高いトップ3の項目名と推定基準値、対応する基準値<sup>6)</sup>およびデータの含有率が示されている。

表2 検査データの推定基準値の例

検査項目名	推定基準値	基準値	データ含有率
WBC (単位: ×10 <sup>3</sup> /μL)	3.09-8.91	3.5-9.0	0.8385
Ht (単位: %)	30.80-46.74	男 40.0-50.0 女 35.0-45.0	0.8385
Hb (単位: g/dL)	9.79-15.72	男 14.0-18.0 女 12.0-16.0	0.8385

観察項目データでは、連続値を対象とするため、離散値と想定される0や1の値の項目は離散化から除外する。観察項目データの離散化の対象は51項目で、表3はデータの含有率の高いトップ3の項目と推定基準値、対応する基準値<sup>7)</sup>およびデータの含有率が示されている。

表3 観察項目データの推定基準値の例

観察項目名	推定基準値	基準値	データ含有率
脈拍数 (単位: /分)	52.97-92.00	45-85	0.9957
収縮期血圧 (単位: mmHg)	91.79-149.21	129 以下	0.9925
拡張期血圧 (単位: mmHg)	49.91-91.06	84 以下	0.9924

### 3.3 算定患者予測と粒度調整の評価

肺血栓塞栓症予防管理料・輸血管管理料 1・麻酔管理料 1 (閉鎖循環式全身麻酔)の算定患者予測に用いるデータは、

宮崎大学医学部附属病院の2015・2016年の2年分の様式1、検査・観察項目データである。患者数は計21,529名である。予測における目的変数は、管理料の算定有無の2値とする。診療データの粒度調整の有効性を検証するため、説明変数は、Kコードの粒度調整の評価の場合にはKコードとする。粒度調整なしの場合は、粒度調整前のKコードのみを用い、粒度調整ありの場合には、粒度調整前のKコードと粒度調整後のKコードを用いる。Kコードの値は1・0(あり・なし)の2値である。検査・観察項目の離散化の評価には、説明変数を検査・観察項目とする。粒度調整なしの場合は、外れ値除去後の粒度調整前の検査・観察項目の元の値を用いる。粒度調整ありの場合には、推定基準値を活用して、推定基準値内の値に対しては重みを0、推定基準値外の異常な値に対する重みを1とする。そして、外れ値除去後の粒度調整前の検査・観察項目の元の値に上述の重みと係数(-5)を掛け合わせた値を用いる。

粒度調整のあり・なしは、5つの分類器(Logistic Regression, Naive Bayes, Neural Network, Random Forest, SVM(linear))による評価指標の平均値を用いて実験的に比較する。本研究では5-foldクロスバリデーションを用いる。分類器のパラメータはデフォルトとする。予測モデルの評価には、AUC、Precision、Recall、F-measureの指標を用いる。3つの管理料ごとで実験を行い、5つの分類器を用いた場合の粒度調整あり・なしの評価指標の平均値を比較することで、粒度調整の有効性を検証する。粒度調整ありの場合の平均値が高かった際には、統計的に有意な差があるかどうかをPaired t-testにより検定する。本実験では、PythonのScikit-learn<sup>8)</sup>を用いる。

#### 4. 結果

表4-9はKコードの粒度調整の有効性を検証するために行った粒度調整あり・なしの実験結果であり、表10、11は3つの管理料の算定予測実験の平均値を示している。表12-17は検査・観察項目の粒度調整の有効性を検証するために行った粒度調整あり・なしの実験結果であり、表18、19は3つの管理料の算定予測実験の平均値を示している。

粒度調整ありの実験結果内では、粒度調整なしと比較して評価値が高い場合に太字で値を示している。その際、Paired t-testを行い、両側確率でp値が0.05未満の場合に、粒度調整あり・なしの評価指標の平均値の差が統計的に有意であることを表内で(< 0.05)と示している。5つの分類器(Logistic Regression, Naive Bayes, Neural Network, Random Forest, SVM(linear))については、表内でそれぞれLR、NB、NN、RF、SVMと表す。表10、11および表18、19では、3つの管理料(肺血栓塞栓症予防管理料、輸血管理料1、麻酔管理料1(閉鎖循環式全身麻酔))をMF1、MF2、MF3と表す。

実験結果から、Kコードの粒度調整の有効性は、Precision以外の評価指標であるAUC、Recall、F-measureで、概して統計的な有意差とともにみられた。特に、F-measureは3つの管理料の算定予測実験の平均で約1%向上した。

検査・観察項目の粒度調整の有効性は、AUCとPrecision以外の評価指標であるRecall、F-measureで、概して統計的な有意差とともにみられた。3つの管理料の算定予測実験の平均でRecallは約6.5%、F-measureは約3%向上した。

表4 肺血栓塞栓症予防管理料の算定予測におけるKコードの粒度調整ありの場合の結果

	AUC	Precision	Recall	F-measure
--	-----	-----------	--------	-----------

LR	<b>0.6595</b>	0.8209	<b>0.2573</b> (< 0.05)	<b>0.3917</b> (< 0.05)
NB	0.6399	0.8469	<b>0.2392</b> (< 0.05)	<b>0.3728</b> (< 0.05)
NN	<b>0.6571</b>	0.7958	<b>0.273</b>	<b>0.4063</b>
RF	<b>0.6569</b> (< 0.05)	0.8095	<b>0.2733</b> (< 0.05)	<b>0.4082</b> (< 0.05)
SVM	<b>0.6588</b> (< 0.05)	0.8107	<b>0.2658</b> (< 0.05)	<b>0.4002</b> (< 0.05)
Average	<b>0.6544</b>	0.8167	<b>0.2617</b>	<b>0.3958</b>

表5 肺血栓塞栓症予防管理料の算定予測におけるKコードの粒度調整なしの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.6559	0.8416	0.2332	0.365
NB	0.6417	0.8895	0.2163	0.3477
NN	0.6556	0.8071	0.2687	0.4026
RF	0.6549	0.8162	0.2703	0.4056
SVM	0.6559	0.8229	0.2595	0.3944
Average	0.6528	0.8355	0.2496	0.3831

表6 輸血管理料1の算定予測におけるKコードの粒度調整ありの場合の結果

	AUC	Precision	Recall	F-measure
LR	<b>0.659</b>	0.7526	<b>0.1606</b> (< 0.05)	<b>0.264</b> (< 0.05)
NB	<b>0.6519</b>	0.6696	<b>0.188</b> (< 0.05)	<b>0.2931</b> (< 0.05)
NN	<b>0.6563</b>	<b>0.7197</b>	<b>0.184</b>	<b>0.2929</b>
RF	0.6543	0.7165	0.1832	0.2912
SVM	<b>0.6577</b>	0.7284	<b>0.1792</b>	<b>0.2873</b>
Average	<b>0.6558</b>	0.7174	<b>0.179</b>	<b>0.2857</b>

表7 輸血管理料1の算定予測におけるKコードの粒度調整なしの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.6571	0.7661	0.1542	0.2561
NB	0.6493	0.7686	0.1625	0.2679
NN	0.6561	0.7195	0.1828	0.2912
RF	0.6545	0.7172	0.184	0.2923
SVM	0.6564	0.7346	0.1788	0.2871
Average	0.6547	0.7412	0.1725	0.2789

表8 麻酔管理料1(閉鎖循環式全身麻酔)の算定予測におけるKコードの粒度調整ありの場合の結果

	AUC	Precision	Recall	F-measure
LR	<b>0.6631</b> (< 0.05)	0.3436	<b>0.0211</b>	<b>0.0391</b>
NB	<b>0.587</b>	<b>0.3513</b>	<b>0.0156</b> (< 0.05)	<b>0.0294</b> (< 0.05)
NN	<b>0.6508</b> (< 0.05)	0.3726	0.0536	0.093
RF	<b>0.6477</b>	<b>0.3537</b>	<b>0.0495</b>	<b>0.0864</b>
SVM	<b>0.6541</b> (< 0.05)	0.3826	<b>0.0449</b>	<b>0.0801</b>
Average	<b>0.6405</b>	<b>0.3608</b>	<b>0.0369</b>	<b>0.0656</b>

表 9 麻酔管理料 1 (閉鎖循環式全身麻酔) の算定予測における K コードの粒度調整なしの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.6506	0.4086	0.0105	0.0198
NB	0.5699	0.1	0.0005	0.0009
NN	0.646	0.3771	0.0564	0.0973
RF	0.6449	0.3532	0.0486	0.085
SVM	0.6478	0.384	0.044	0.0787
Average	0.6318	0.3246	0.032	0.0563

表 10 3 つの管理料の算定予測における K コードの粒度調整ありの場合の結果の比較

	AUC	Precision	Recall	F-measure
MF1	0.6544	0.8167	0.2617	0.3958
MF2	0.6558	0.7174	0.179	0.2857
MF3	0.6405	0.3608	0.0369	0.0656
Average	<b>0.6503</b>	0.6316	<b>0.1592</b>	<b>0.249</b>

表 11 3 つの管理料の算定予測における K コードの粒度調整なしの場合の結果の比較

	AUC	Precision	Recall	F-measure
MF1	0.6528	0.8355	0.2496	0.3831
MF2	0.6547	0.7412	0.1725	0.2789
MF3	0.6318	0.3246	0.032	0.0563
Average	0.6464	0.6338	0.1514	0.2394

表 12 肺血栓塞栓症予防管理料の算定予測における検査・観察項目の粒度調整ありの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.5036	0	0	0
NB	0.5048	<b>0.2754</b>	<b>0.496</b> ( $< 0.05$ )	<b>0.3538</b>
NN	0.4976	<b>0.2669</b>	0.1887	0.2135
RF	0.4985	0.2636	<b>0.062</b> ( $< 0.05$ )	<b>0.1003</b> ( $< 0.05$ )
SVM	<b>0.5</b>	0.2734	<b>0.4191</b>	<b>0.3269</b>
Average	0.5009	0.2159	<b>0.2332</b>	<b>0.1989</b>

表 13 肺血栓塞栓症予防管理料の算定予測における検査・観察項目の粒度調整なしの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.5074	0	0	0
NB	0.5078	0.272	0.2879	0.2703
NN	0.5002	0.2633	0.5242	0.321
RF	0.5074	0.2772	0.0458	0.0785
SVM	0.4993	0.2745	0.2908	0.2298
Average	0.5044	0.2174	0.2297	0.1799

表 14 輸血管管理料 1 の算定予測における検査・観察項目の粒度調整ありの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.4912	0	0	0
NB	0.5035	0.1179	<b>0.4392</b> ( $< 0.05$ )	<b>0.1858</b> ( $< 0.05$ )
NN	0.5018	<b>0.1204</b>	<b>0.1268</b>	<b>0.114</b>
RF	<b>0.4991</b>	0.0651	0.0032	0.0061

SVM	0.4913	<b>0.1108</b>	<b>0.1741</b> ( $< 0.05$ )	<b>0.1332</b> ( $< 0.05$ )
Average	0.4974	0.0828	<b>0.1487</b>	<b>0.0878</b>

表 15 輸血管管理料 1 の算定予測における検査・観察項目の粒度調整なしの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.5129	0	0	0
NB	0.5125	0.1202	0.1395	0.1284
NN	0.5062	0.1202	0.0437	0.0519
RF	0.4952	0.2241	0.0048	0.0093
SVM	0.4941	0.0832	0.0465	0.0467
Average	0.5042	0.1095	0.0469	0.0472

表 16 麻酔管理料 1 (閉鎖循環式全身麻酔) の算定予測における検査・観察項目の粒度調整ありの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.5132	0	0	0
NB	0.5099	0.103	<b>0.4773</b> ( $< 0.05$ )	<b>0.1693</b>
NN	0.4965	0.1014	<b>0.0852</b> ( $< 0.05$ )	<b>0.0885</b> ( $< 0.05$ )
RF	0.5021	<b>0.1038</b> ( $< 0.05$ )	<b>0.0041</b> ( $< 0.05$ )	<b>0.0079</b> ( $< 0.05$ )
SVM	<b>0.5004</b>	<b>0.0988</b>	<b>0.2836</b>	<b>0.1453</b>
Average	0.5044	<b>0.0814</b>	<b>0.17</b>	<b>0.0822</b>

表 17 麻酔管理料 1 (閉鎖循環式全身麻酔) の算定予測における検査・観察項目の粒度調整なしの場合の結果

	AUC	Precision	Recall	F-measure
LR	0.5147	0	0	0
NB	0.521	0.1083	0.268	0.1532
NN	0.502	0.1454	0.0128	0.022
RF	0.5065	0	0	0
SVM	0.496	0.0917	0.1062	0.0531
Average	0.508	0.0691	0.0774	0.0457

表 18 3 つの管理料の算定予測における検査・観察項目の粒度調整ありの場合の結果の比較

	AUC	Precision	Recall	F-measure
MF1	0.5009	0.2159	0.2332	0.1989
MF2	0.4974	0.0828	0.1487	0.0878
MF3	0.5044	0.0814	0.17	0.0822
Average	0.5009	0.1267	<b>0.1839</b>	<b>0.123</b>

表 19 3 つの管理料の算定予測における検査・観察項目の粒度調整なしの場合の結果の比較

	AUC	Precision	Recall	F-measure
MF1	0.5044	0.2174	0.2297	0.1799
MF2	0.5042	0.1095	0.0469	0.0472
MF3	0.508	0.0691	0.0774	0.0457
Average	0.5055	0.132	0.118	0.0909

## 5. 考察

5 つの分類器を用いた 3 つの管理料の算定予測の実験結

果から、K コードと検査・観察項目の粒度調整の有効性が Recall と F-measure の評価指標でみられたことから、Precision と Recall の調和平均である F-measure の高さにより総合的に管理料算定患者の予測に有効であること、また、Recall の高さにより網羅的に管理料を算定した患者の中で管理料を算定したと予測する際に有効であることが示唆される。本実験では 5 つの分類器を用いた評価指標の平均値を比較していることから、粒度調整は予測問題において分類器に依存しないことが示唆される。

本研究で対象とした K コード以外の診療概念コードの粒度調整を行う際は、粒度調整の手法として階層クラスタリングは活用できるが、分析目的に応じてどのような視点で粒度を調整するか、つまり、階層クラスタリングの要素をどのように設定するかが重要になると考えられる。検査・観察項目の粒度調整では、入院患者の検査・観察項目の正常値・異常値を判別するために正規分布を活用していることから、検査・観察項目のデータが少ない場合には、判別が困難であることが考えられる。しかしながら、公開されている検査・観察項目の基準値の単位が、実際に用いるデータと異なる場合に加工が不必要なこと、また公開されていない検査・観察項目がある場合においても正常・異常値に判別できることから、効率的に検査・観察項目の粒度調整が可能であると考えられる。

今後の課題は、予測の目的変数が管理料以外の場合における粒度調整の影響を分析することである。また、K コード以外の診療概念コードの粒度調整を行い、その有効性を検証する。そして、診療データの粒度調整という前処理の良し悪しを蓄積するための機能を構築する。

膨大な医療情報が利活用可能な中、診療概念コードや検査・観察項目のデータを臨床研究や多変量解析などの分析の目的に適した形に加工することは、医療情報の二次利用の促進に寄与することが期待される。粒度調整の失敗結果も試行錯誤のプロセスとしてノウハウを蓄積できる仕組みがあれば、さらなる二次利用の促進が期待できると考えられる。

## 6. 結論

本研究は、後向き分析を効率よく進めるための医療情報の利活用基盤として、臨床研究における情報処理フレームワークを提案し、フレームワーク内の診療データの粒度調整の有効性を検証した。3 つの管理料の算定予測の実験結果から、粒度調整あり・なしについて、5-fold クロスバリデーションを用いた 4 つの評価指標の 5 つの分類器による平均値を比較すると、粒度調整をすることで概して Recall と F-measure (Precision と Recall の調和平均) の評価値が改善する傾向がみられた。このことから、総合的に管理料算定患者の予測に有効であり、網羅的に管理料算定患者の中で管理料を算定したと予測する際に有効であることが示唆された。また、5 つの分類器で評価指標の平均値を比較していることから、粒度調整は予測問題において分類器に依存しないことが示唆された。診療概念コードや検査・観察項目といった診療データの粒度を効果的に調整できることは、過去の医療情報リアルワールドデータを用いた後向き分析の効率化に向けて期待が持たれる。今後は、考察で述べた課題に取り組むとともに、群間比較の分析フェーズにおいての方法論の確立を目指す。

## 謝辞

本研究は JSPS 科研費 JP18K09948 の助成を受けたもの

です。

## 参考文献

- 1) 健康・医療戦略推進本部. 次世代医療基盤法の施行について, 2018.  
[[http://www.kantei.go.jp/jp/singi/kenkouiryou/jisedai\\_kiban/houritsu.html](http://www.kantei.go.jp/jp/singi/kenkouiryou/jisedai_kiban/houritsu.html) (cited 2018-Aug-23)].
- 2) Lorenzi, EC., Brown, SL., Sun, Z., Heller, K. Predictive Hierarchical Clustering: Learning clusters of CPT codes for improving surgical outcomes. In Machine Learning for Healthcare Conference 2017 ; 231-242.
- 3) Lustgarten, JL., Gopalakrishnan, V., Grover, H., Visweswaran, S. Improving classification performance with discretization on biomedical datasets. In AMIA annual symposium proceedings 2008 ; 2008 : 445-449.
- 4) Maslove, DM., Podchyska, T., Lowe, HJ. Discretization of continuous features in clinical datasets. Journal of the American Medical Informatics Association 2012 ; 20(3) : 544-553.
- 5) Jones E, Oliphant E, Peterson P. SciPy: Open Source Scientific Tools for Python, 2001.  
[<http://www.scipy.org/> (cited 2018-Aug-23)].
- 6) 日本臨床検査医学会ガイドライン作成委員会. 付録: 基準範囲、パニック値/緊急報告値. 臨床検査のガイドライン JSLM2012, 日本臨床検査医学会, 2012 : 391-396.
- 7) 公益社団法人日本人間ドック学会. 平成 30 年度 (2018 年度) からの基本検査項目表・判定区分表について, 2018.  
[<https://www.ningen-dock.jp/other/inspection> (cited 2018-Aug-23)].
- 8) Pedregosa, F., Varoquaux, G., Gramfort, A., et al. Scikit-learn: Machine learning in Python. Journal of machine learning research 2011 ; 12(Oct) : 2825-2830.

