

厚生労働科学研究費補助金（政策科学総合研究事業（政策科学推進研究事業））
分担研究報告書

横断的特徴選択に基づく介護レセプトデータの解析

研究代表者 田宮菜奈子 筑波大学医学医療系ヘルスサービスリサーチ分野 教授
研究協力者 佐久間淳 筑波大学システム情報系情報工学域 准教授

研究要旨

近年、多数の特徴量を持つ時系列データを収集することができるようになり、こうしたデータから意義のある特徴量を発見する様々な手法が発展した。我々は、スパース正則化項を用いたマルチタスク学習によりスパース推定を用いて、介護サービスの利用回数や背景情報、要介護度を記録した時系列データである介護レセプトデータに対して分析を行い、時間横断的な要介護度と関連の強い介護サービスを抽出し、その関連度の強さを定量化した。

A．研究目的

医療や介護の分野で人のサービス利用履歴や状態を表す指標が時系列データとして蓄積されている。しかし、こうした時系列データは膨大な量となりやすく、分析者の知識に頼った分析では有用な情報を見落としやすい。これはデータの量が多いためだけでなく、入力データが特徴量 \times サンプル数 \times 時間、出力データが評価指標 \times サンプル数 \times 時間というような高次テンソルとなるためデータを俯瞰的にとらえることが難しいためである。

こうした時系列データに対する伝統的な分析手法としては時系列性を無視した重回帰分析や二つの時刻点間の変化に対する重回帰分析が行われてきた。しかし、前者は本来存在するはずの時間変化を考慮していないという問題があり、後者はデータに対して適切な仮説をおくことができるような十分な知識が必要となる。

この研究では、各月の介護サービスの利用回数と要介護度、要介護者の背景情報を含む時系列データに対して、十分な知識を用いることなくデータ自身の特性を発見することが目的である。より具体的には、介護

レセプトデータに対して時間横断的な特徴選択を行うことで、どのような介護サービスが要介護度の高低に関連し、そうした関連性が時間変化によってどう変化するのかを発見する。

B．研究方法

機械学習を用いた時系列データに対する分析手法として、スパース正則化項を用いたマルチタスク学習による時間横断的な特徴選択の手法を提案した。

介護レセプトデータのように時系列データに欠損値を含む場合には、そうした欠損値を含むデータを取り除く必要があるため、既存手法では多数のサンプルを確保することが困難であった。

我々はサンプルに欠損値が含まれている場合にサンプルの欠損値を含んでいる時刻点のみを取り除き、隣接時刻点を考慮して欠損した時刻点を補完しモデルを構築する方法を提案した。これにより、欠損値を取り除いたときより、より高い精度で学習できるようになった。

得られた学習モデルの係数は、要介護度の高さ・低さと各種サービスの関連の強さと

その時間変化を表す。これらの係数によって、要介護度の高さ・低さと各種サービスの関連性を分析した。

(倫理面への配慮)

本研究で用いた個人情報保護法
の遵守を徹底の上、研究倫理委員会の承認
のもと研究を実施した。

C．研究結果

提案法を用いた横断的特徴選択の結果、以下
の結果を得た。

- (1) 「介護福祉施設サービス」は高い要介護
度強く関連していた。これは介護の負
担が大きい人が施設に入るためである
と考えられる。
- (2) 「訪問介護」、「訪問入浴」、「訪問看
護」のような一般的なサービスが高い
要介護度と関連することがわかった
- (3) 「福祉用具貸与」が高い要介護度と関
連が大きいことがわかった
- (4) 予防に関連するサービスは一般的に状
態が良い人が受けるサービスであるが、
このようなサービスに対してもリハビリ
や通所介護は低い要介護度と関連が
大きいことがわかった
- (5) 年齢は高いことは高い要介護度と関連
が大きいことがわかった
- (6) 「訪問リハビリ」、「通所介護」、「通
所リハビリ」を受けている人は受けて
いない人に比べて要介護度が低い傾向
があり、受け始めたときより受けてか
ら時間が経過したほうが低い要介護度
と関連が大きいことがわかった

D．考察

横断的特徴選択によって、要介護度の高低
と介護サービスの関連の強さを定量的に可
視化することができた。「訪問リハビリ」
を受けている人が介護サービスを受ける期
間が長くなるほど、低い要介護度との関連

性が高まるという考察を得た。要介護度の
改善を目的とする「訪問リハビリ」につい
てこのような結果を得たことは、こうした
サービスの重要性を示す有意義な結果であ
った。

E．結論

分析結果の多くは専門家の知識と一致し、
提案手法によりデータを俯瞰的に分析でき
ることがわかった。これまでの介護サービ
スの研究では介護サービス種類コードに対
して分析が行われていたが、本研究の目的
関数では大きなデータに対しても分析する
ことができるため、介護サービス項目コー
ドを用いた分析も検討したい。

F．研究発表

1．論文発表
なし

2．学会発表

中里佳央, 佐久間淳, 川村顕, 田宮菜緒子,
介護レセプトデータに対する横断的特徴選
択による介護サービスの評価, 第 19 回情
報論的学習理論ワークショップ, D-74,
2015 年 11 月 27 日.

G．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