論理制約を考慮したテーブルデータを対象とした予測モデル構築フレームワーク

A framework to construct predictive models with logical constraints for table data

尾上 圭介 [*1]　小島 諒介 [*1]
Keisuke Onoue　Ryosuke Kojima

[*1]京都大学 医学研究科
Graduate School of Medicine, Kyoto University

Reliability of machine learning models are getting serious in the many application areas such as medical and business fields. One approach to addressing these requirements is to use logical constraints representing background knowledge to prevent the model from producing outputs that violate the constraints. However, this approach requires manual setting of all logical constraints for the target task, which is very labor intensive. In this study, we propose a framework that combines RuleFit, a machine learning-based method for automatically acquiring rules, and a method for building predictive models under logical constraints for table data. We evaluate our proposed framework by the prediction accuracy and the violation rate of the constraints using a diabetes benchmark dataset. Using our proposed framework, we achieved to identify the best method from the viewpoint of prediction accuracy and constraint violation rate.

1. はじめに

医療分野やビジネス分野などの分野では，機械学習の活用において，モデルの予測性能のみならず信頼性が重視される場合がしばしばあり，そのアプローチの一つとして，機械学習モデルの出力を論理制約によって制限することで望ましい出力とする方法が期待されている．これらを実現するために機械学習に制約充足問題や制約付きの最適化問題を取り込む方法が提案されており [Goyal 22]，他にも類似のアプローチとして，深層学習と記号処理の融合の一部としても取り組まれている [Kautz 22]．具体的には，これらのアプローチでは深層学習の層に論理制約を組み込む方法 [Wang 23] や損失関数に制約を組み込む方法 [Yang 22, Roychowdhury 21] など多数の手法が近年提案されている．

そのような背景のもと，Giannini らは，背景知識を多値論理の一つであるŁukasiewicz 論理の論理式として表現し，それに対応する不等式を制約としてカーネル法の学習スキームに組み込む方法を提案している [Giannini 17]．この手法は，カーネル法との相性が良い点や二次計画問題として定式化できるといった利点から，データ数が少なく深層学習が難しい場面でも安定して動作するという利点が期待できる．しかし，この手法を適用するには，全ての論理式を手動で設定する必要から制約の作成に非常に労力がかかるという問題が存在する．

そこで，本研究ではテーブルデータを対象にし，ルールの自動的な獲得手法と論理制約下での予測モデルの構築手法を組み合わせることによって，ルール作成の手間を省力化した制約付き予測モデル構築フレームワークを提案する．また，ルール獲得を自動化することで人手によるルール作成の伴わない客観的な制約付きの機械学習モデルの評価が可能になることも期待できる．本フレームワークを用いた評価実験として，糖尿病のベンチマークデータセット [Smith 88] を用いて予測性能とルールの充足率を評価した．本フレームワークの実装および実験データの詳細は GitHub リポジトリ [*1] より入手できる．

連絡先: 小島諒介, 京都大学医学研究科, 京都府京都市左京区聖護院川原町 54
E-mail:kojima.ryosuke.8e@kyoto-u.ac.jp

2. 提案手法

2.1 概要

図 1 に提案フレームワークの概略図を示す．本フレームワークではテーブルデータを対象とし，データからルールの自動抽出を行い，得られたルールセットから各ルールを論理制約として表現し，これらの制約下での予測モデルの学習を行う．以下では，これらの流れに沿って各手順の詳細を述べる．

2.2 ルールの自動抽出

テーブルデータからのルールの自動抽出のためのアルゴリズムには RuleFit [Friedman 08] を採用した．RuleFit を適用する前の前処理として，テーブルデータの各列ごとに離散化と各特徴量の One-Hot encoding を行う．例えば，与えられたテーブルデータ (図 1 中の (a)) の BMI 列を「低 (L)，中 (M)，高 (H)」の 3 段階に離散化し，One-Hot Encoding することによって新たなテーブル（1 中の表 (b)）を得る．この操作を正解ラベルである Outcome 列以外の全ての列に対して行う．

この変換後のテーブルデータに対して RuleFit を適用し，図 1 中の表 (c) のようなルールセットを生成する．RuleFit は特徴量をもとにルールを作成し，そのルールに重みを付けた線形モデルで予測器を構築する手法である．ここでのルールには，例えば「ある特徴がある値以上である」といった決定規則も含まれており，これらを連言（$\wedge$）で組み合わせた規則も作ることができる．上記をまとめてルールセットと呼ぶことにする．

2.3 制約論理式の構築

変換後のテーブルデータ (表 (b)) の各特徴量を P_j ($j \in \mathbb{N}_J$) とする．これ以降簡単のために，$\mathbb{N}_J = \{1, \ldots, J\}$ と書くことにする．

表 (c) の形のルールセットにおいて，ルールの条件部分の記述は「P_j」,「不等号」,「条件の連言（条件を全て満たす）」 の組み合わせによって構成されているものとする．

サンプル x の各特徴量の値は述語 $p_j(\mathbf{x})$，$j \in \mathbb{N}_J$ と対応付けて，表 1 の規則に従って実数上の関数に変換する [Giannini 17]

*1 `https://github.com/k-onoue/lukasiewicz_2`

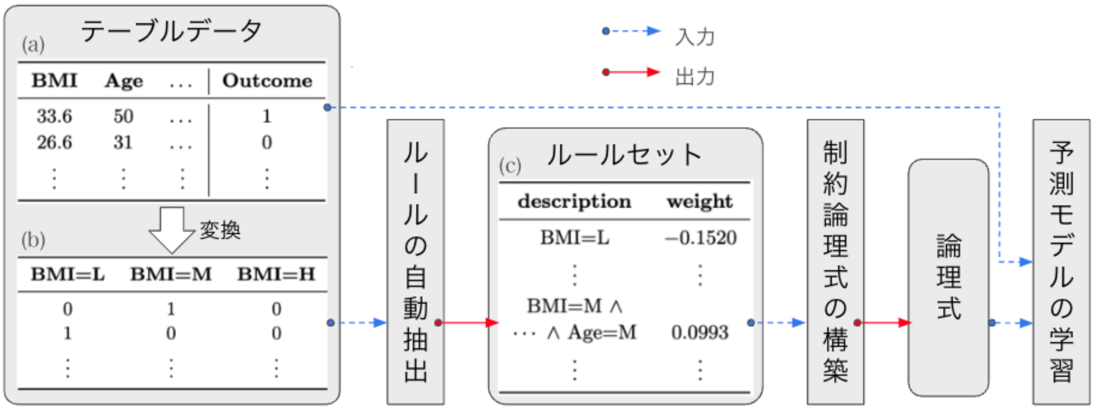

図 1: 提案フレームワークの流れ図：角丸四角形は入出力を，長方形は入力に対する処理を表す

変換前	変換後
$P_j \wedge P_{j'}$	$p_j(\mathbf{x}) \otimes p_{j'}(\mathbf{x})$
$P_j = 1$	$p_j(\mathbf{x})$
$P_j = 0$	$\neg\, p_j(\mathbf{x})$

表 1: 条件部分の変換規則

*2．ここでは，論理積 $\otimes$ は $x \otimes y = \max\{0, x+y-1\}$ で定義する．例えば，Pregnancies_M $= 0 \wedge \cdots \wedge$ BMI_M $= 1$ は $\neg$Pregnancies_M$(\mathbf{x}) \otimes \cdots \otimes$ BMI_M$(\mathbf{x})$ のように変換する．

正解ラベルに相当するテーブルデータの列名を P' とすると，weight の値が正ならば $p'(\mathbf{x})$，負ならば $\neg\, p'(\mathbf{x})$ とし，最後に変換後の規則と正解ラベルを含意 $\rightarrow$ で繋ぐことによってルールから論理式への変換とする．ただし含意は，$x \rightarrow y = \min\{1, 1-x+y\}$ と定義する．例えば，(Pregnancies_M $\leq 0.5 \wedge \cdots \wedge$ BMI_M > 0.5, 0.0993) は $\neg$Pregnancies_M$(\mathbf{x}) \otimes \cdots \otimes$ BMI_M$(\mathbf{x}) \rightarrow$ Outcome$(\mathbf{x})$ となる．

2.4 予測モデルの学習

学習には元のテーブルデータと RuleFit の適用の際に離散化して得られたテーブルデータからなる教師ありデータ $\mathscr{L} = \{\mathscr{L}_j : j \in \mathbb{N}_j\}$ 述語ごとの教師ありデータは，$\mathscr{L}_j = \{(\mathbf{x}_l, y_l^j) : l \in \mathbb{N}_L\}$ とあらわされる．ここで，$\mathbf{x}_l \in \mathbb{R}^n, y_l^j \in \{-1, 1\}$ であり，n は特徴量の数で，L はサンプル数である．

教師なしデータ $\mathscr{U} = \{\mathscr{U}_j : j \in \mathbb{N}_J\}$，を使用する．ただし，$\mathscr{U}_j = \{\mathbf{x}_u : u \in \mathbb{N}_U, \mathbf{x}_u \sim \text{Uniform}\}$ とする．教師なしデータ x_u は教師データをもとにした各特徴量の定義域から一様に U 個サンプリングする．

述語 p_j を各データ点 $\mathbf{x}$ に対して計算するための具体的な関数表現を与える．各 j に対して特徴空間へのマップ $\phi_j : \mathbb{R}^n \rightarrow \mathbb{R}^{N_j}$ を用いて，$p_j : \mathbb{R}^n \rightarrow \mathbb{R}$ は以下のように表現される．

$$p_j(\mathbf{x}) = \omega_j \cdot \phi_j(\mathbf{x}) + b_j$$

ただし，$\omega_j \in \mathbb{R}^{N_j}$，$b_j \in \mathbb{R}$ は学習する重みとし，ルールの数を H とする．

各 p_j と対応する教師なしデータ $\mathscr{U}_j$ について，$\mathscr{U}_j$ の全ての元 $\mathbf{x}_u$ での p_j の評価をまとめて $\bar{\mathbf{p}}_j = (p_j(\mathbf{x}_1), \ldots, p_j(\mathbf{x}_U))$ と表し，さらに $\bar{\mathbf{p}} = (\bar{\mathbf{p}}_j, \ldots \bar{\mathbf{p}}_J)$ とする．

ここで，h 番目 $(h \in \mathbb{N}_H)$ のルールから前節で述べた制約論理式によって与えられる関数 $f_h : [0,1]^{J \times U} \rightarrow [0,1]$ を考える．f_h は教師なしのデータ $\bar{\mathbf{p}}$ を引数に取ることで，h 番目のルールの $\mathscr{U}$ 上での評価を意味する．例えば，$f_h(\bar{\mathbf{p}}) = 1$ は $\mathscr{U}$ が h 番目のルールを充足することを意味する．

以上の設定を踏まえて，以下の制約付き最小化問題を二次計画法を用いて解く．

$$\min \frac{1}{2} \sum_{j=1}^{J} \|\omega_j\|^2 + C_1 \sum_{j=1}^{J} \sum_{l=1}^{L} \xi_{j_l} + C_2 \sum_{h=1}^{H} \xi_h \tag{1}$$

subject to

$$y_l^j (2p_j(\mathbf{x}_l) - 1) \geq 1 - 2\xi_{j_l}, \quad \xi_{j_l} \geq 0, \tag{2}$$

$$1 - f_h(\bar{\mathbf{p}}) \leq \xi_h, \quad \xi_h \geq 0, \tag{3}$$

$$0 \leq p_j(\mathbf{x}_s) \leq 1. \tag{4}$$

ただし，$j \in \mathbb{N}_j, l \in \mathbb{N}_L, (\mathbf{x}_l, y_l^j) \in \mathscr{L}_j, h \in \mathbb{N}_h, s \in \mathbb{N}_{L+U}, \mathbf{x}_s \in \mathscr{U}_j \cup \{\mathbf{x}_l : (\mathbf{x}_l, y_l^j) \in \mathscr{L}_j\}$ で，ξ_{j_l}, ξ_h はスラック変数である．

不等式 (2) は各分類器 p_j のソフト制約で，パラメータ C_1 がその制約の厳しさを表す．論理制約は (3) で表現されており，C_2 は論理制約の厳しさを表す．さらに不等式 (4) は述語 p_j の $\mathscr{L}_j$，$\mathscr{U}_j$ 上での評価値が $[0,1]$ の範囲に収まることを示している．また上記の主問題に対して通常のカーネル法と同様に双対問題を代わりに解くこともできる．

3. 実験

本実験では後述の糖尿病のベンチマークデータセットに対して，提案するフレームワークによる予測モデルを構築し，予測精度と制約の充足率を調べる．また，既存の機械学習モデルとの比較を行う．

3.1 実験設定

本実験で用いるデータセットは Pima Indian Diabetes [Smith 88] である．このデータセットは糖尿病の医療検査・診断に関するデータセットで，BMI，年齢，インスリンレベル等の８つの検査項目の測定値と実際に糖尿病であるかどうかの診断結果 (Outcome) から構成される．

*2 この変換はLukasiewicz 論理に由来しており，本稿では簡単のために論理式とそれがあらわす関数が混同の恐れがない限り同一視して扱う．

モデル	AUC	違反率
l-SVM	0.838 ± 0.045	0.880 ± 0.029
r-SVM	0.838 ± 0.049	0.908 ± 0.039
LogReg	0.841 ± 0.047	0.886 ± 0.026
RuleFit-d	0.816 ± 0.052	0.805 ± 0.084
RuleFit	0.795 ± 0.033	0.881 ± 0.065
RF-d	0.772 ± 0.049	0.739 ± 0.068
RF	0.833 ± 0.040	0.880 ± 0.047
l-SVM-p	0.537 ± 0.318	0.630 ± 0.049
r-SVM-p	0.823 ± 0.046	0.664 ± 0.087
LogReg-p	0.827 ± 0.051	0.552 ± 0.110

表 2: 実験結果 (総ルール数: 34.8 ± 3.2)

予測モデルの評価には AUC (Area Under the ROC Curve) と「ルールの充足率」を用いる．制約の充足率とは，与えられたルールセットうち，テストデータの任意の点で充足したルールの割合である．また 1 − 充足率 をルール違反率として定義する．

以下の予測モデルを評価する．

1. SVM (線形カーネル) [l-SVM]
2. SVM (RBF カーネル) [r-SVM]
3. ロジスティック回帰 [LogReg]
4. RuleFit (離散) [RuleFit-d]
5. RuleFit [RuleFit]
6. ランダムフォレスト (RuleFit-d 内の予測器) [RF-d]
7. ランダムフォレスト (RuleFit 内の予測器) [RF]
8. SVM (線形カーネル) (提案手法) [l-SVM-p]
9. SVM (RBF カーネル) (提案手法) [r-SVM-p]
10. ロジスティック回帰 (提案手法) [LogReg-p]

[ ] はこれ以降用いる略称とする．基本的には入力の特徴量は連続値を用いるが，RuleFit についてはルール生成の際に離散化したデータ上 (RuleFit，RF) でも評価をする．

3.2 実験結果

各予測器の予測性能とルール違反率を表 2 に示した．この結果から従来手法は全て違反率 70 %以上であったのに対し，提案手法ではこれらを下回った結果を出すことに成功した．また，提案手法の中でも，l-SVM-p は AUC は低いが，充足率は良い結果となり，r-SVM-p は AUC，充足率ともに高くなる傾向にあった．LogReg-p は両者を超える性能であった．この結果より，学習時の論理制約の有無は，予測モデルのルールの充足率に影響を与えていることが確認できた．

次に提案手法の性質を調べるために，論理制約の重みパラメータ C_2 の値を変化させる実験と教師なしデータの点数を変える実験を行った．

その結果，C_2 が大きくなるに従って，予測精度には大きな変化は見られなかったが，ルール違反率は下がる傾向が見られた (図 2)．

教師なしデータの点数については予測精度，ルール違反率ともに大きな変化が見られなかった (図 3)．教師なしデータのサイズと計算時間はトレードオフの関係にあるため，調整が必要である．

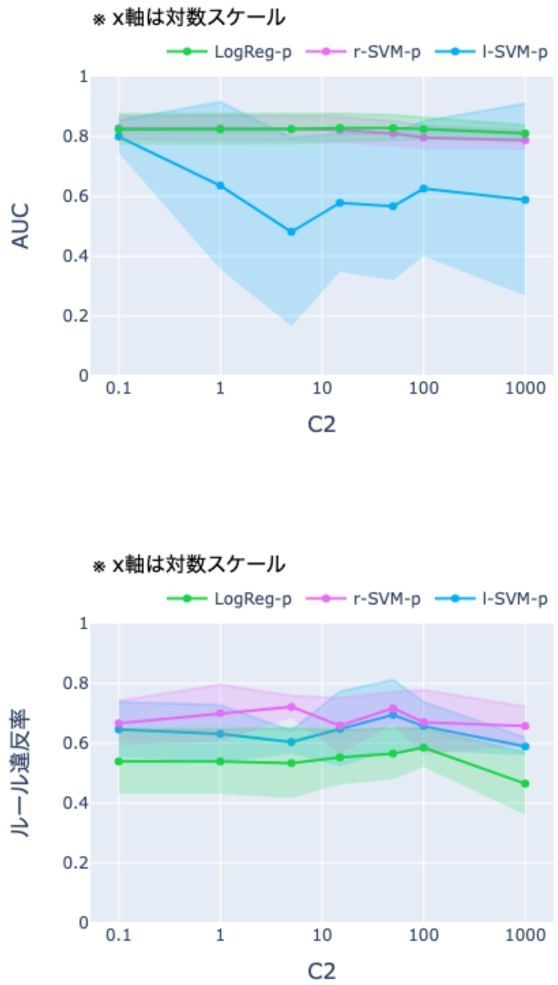

図 2: C_2 を変更した場合の AUC とルール違反率

4. おわりに

本研究では，テーブルデータを対象にしたルールの自動獲得とそれを基にした論理制約下での予測モデル構築を一連の手続きとしたフレームワークを提案した．糖尿病のベンチマークで提案フレームワークを評価し，既存の機械学習モデルに比べ提案手法による予測モデルは，予測性能とルールの充足率において高い性能を示すことが分かった．今後の課題としては，自動獲得したルールの評価や現実問題へ適用した際のルールの有用性の確認といったことが挙げられる．

謝辞

本研究は JSPS 科研費 No. 19KK0260, 20H00475, 21H04905 および CREST JPMJCR22D3 の助成を受けた．

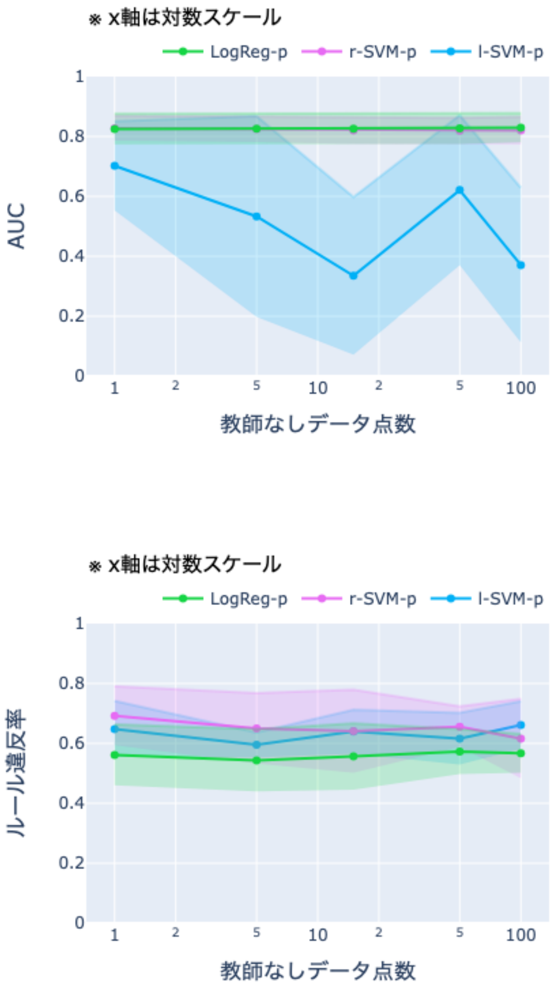

図 3: 教師なしデータ点数 U を変更した場合の AUC とルール違反率

参考文献

[Friedman 08] Friedman, J. H. and Popescu, B. E.: Predictive learning via rule ensembles (2008)

[Giannini 17] Giannini, F., Diligenti, M., Gori, M., and Maggini, M.: *Learning Łukasiewicz Logic Fragments by Quadratic Programming*, pp. 410–426 (2017)

[Goyal 22] Goyal, K., Dumancic, S., and Blockeel, H.: Sade: Learning models that provably satisfy domain constraints, in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 410–425 (2022)

[Kautz 22] Kautz, H.: The third ai summer: Aaai robert s. engelmore memorial lecture, *AI Magazine*, Vol. 43, No. 1, pp. 105–125 (2022)

[Roychowdhury 21] Roychowdhury, S., Diligenti, M., and Gori, M.: Regularizing deep networks with prior knowledge: A constraint-based approach, *Knowledge-Based Systems*, Vol. 222, p. 106989 (2021)

[Smith 88] Smith, J. W., Everhart, J. E., Dickson, W., Knowler, W. C., and Johannes, R. S.: Using the ADAP learning algorithm to forecast the onset of diabetes mellitus, in *Proceedings of the annual symposium on computer application in medical care*, p. 261 (1988)

[Wang 23] Wang, Z., Vijayakumar, S., Lu, K., Ganesh, V., Jha, S., and Fredrikson, M.: Grounding Neural Inference with Satisfiability Modulo Theories, in *NeurIPS* (2023)

[Yang 22] Yang, Z., Lee, J., and Park, C.: Injecting logical constraints into neural networks via straight-through estimators, in *ICML*, pp. 25096–25122PMLR (2022)