

# ブートストラップ法を用いた Alternating Decision Forests の 適応的な汎化性能向上に関する研究

情報数理応用研究

5214C035-6  
指導教員

三沢翔太郎  
後藤正幸

## Improving Generalization Ability of Alternating Decision Forests with Bootstrap Method

MISAWA Shotaro

### 1 研究背景・目的

経営工学の分野では、対象とする問題にあわせた様々な機械学習手法が活用されている。本研究ではこれらのうち、与えられたデータの離散カテゴリラベルを予測する問題にあわせた機械学習手法に着目する。この問題は、説明変数群とカテゴリラベルで構成された学習データ集合が与えられたもとで、新規データのカテゴリを予測する問題である。本研究では、この問題を解く手法の中で、モデル理解の容易さや計算時間の短さなどで有用性がある決定木 [1] に着目する。決定木はノードと分岐条件からなる木構造のモデルであり、これをアンサンブルすることで、より高性能な予測が行えることが知られている。決定木をアンサンブルする手法は Random Forests [2] などが有名であるが、本研究では特に予測精度の高い Alternating Decision Forests (ADF) [3] に着目する。

ADF は各学習データへのウェイトの付与と、深さの同じノードに対する分岐条件の付与を繰り返すことで決定木集合を生成する手法である。また、各ノードへの分岐条件付与において、分岐条件の候補となる説明変数の一部に限定する。これにより、各木が過度に学習データへ当てはまることを避けることで、決定木間の相関を下げていく。しかし、この手法には以下の2つの問題点が存在する。1つ目は外れ値が混入した場合、その影響を強く受け、決定木集合に過学習が生じる点である。2つ目は、ノードごとに分岐条件の候補を変化させることで、学習データへの当てはまりが過度に抑制され、全体の予測精度が低下している可能性がある点である。

そこで本研究では、以下の2つの方法で予測精度の向上を図る。1つ目はウェイト算出時に学習データの外れ値の度合いを考慮することにより、過学習を防止する。2つ目は同一の木にあるノードに対しては同じ説明変数群を繰り返し分岐条件の候補とすることで、割り当てた説明変数群に対する学習データへの当てはまりをよくする。一方で、割り当てた説明変数を用いて予測に有効な分岐条件が付与されなくなった際に、割り当てた説明変数を変更し、過学習や決定木の表現能力の低下を防ぐ。ここで、本研究では外れ値の度合いの推定と説明変数を割り当て直す基準を求めるため、各決定木に対して学習に用いないデータ (OOB データ) を割り当て、各木がそれらのデータを予測した結果を活用する。また、UCI 機械学習レポジトリ [4] のデータと人工的に外れ値を付与したデータを用いた実験により提案手法の有効性を示す。

### 2 準備

#### 2.1 問題設定

$N$  個の学習データの集合を  $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ 、離散カテゴリ集合を  $\mathcal{C} = \{1, \dots, K\}$  とする。  $\mathbf{x}_i$  は  $q$  次元の説明変数ベクトルで、対応する目的変数  $y_i \in \mathcal{C}$  は  $\mathbf{x}_i$  に付与されているカテゴリとする。本研究では、  $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$  が得られたもとで、新たにカテゴリラベルが未知である  $\mathbf{x}_{N+1}$

が与えられたとき、対応する目的変数  $y_{N+1}$  を予測する問題を考える。

#### 2.2 決定木

決定木は図1のようなノードと枝から成る木構造のモデルである。ノードは決定節と葉ノードがあり、決定節はデータ集合をある説明変数を基準に分割する分岐条件を所持するノードで、葉ノードは決定木の末端のノードである。

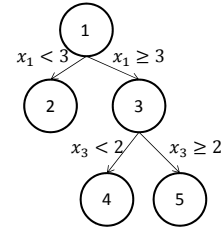


図1. 決定木の例

決定木を生成する際は、根ノードから順に学習データ集合の分割によって、各カテゴリの学習データの割合がなるべく偏るような分岐条件をノードに付与することを繰り返す。また、構築された決定木から新規入力データのカテゴリ予測を行う際は、到達した葉ノードに存在する学習データの所属カテゴリの多数決など、葉ノードに付与された予測規則に従って予測する。

#### 2.3 決定木のアンサンブル

決定木は表現能力が高いため、汎化性能が低くなり易いという問題点がある。これを解決する方法として決定木のアンサンブル手法が広く知られており、その有用性が示されている [2], [3]。アンサンブル手法は複数の異なる性質の決定木を生成し、個々の結果から多数決を行うことで予測を行う手法であり、各決定木の精度を下げることで過学習を抑えつつ決定木間の相関を下げることで個々の低い予測結果を補完し、森全体として汎化性能を向上させている。

### 3 Alternating Decision Forests

本研究では、決定木のアンサンブル手法のうち特に予測性能が優れている ADF に着目する。ADF は、学習データへのウェイトの付与とノードへの分岐条件の付与を深さ毎に繰り返すことにより決定木集合を生成する手法である。これにより、効率よく予測結果の補完を行い、一般的な決定木のアンサンブル手法より高い汎化性能を得る。

#### 3.1 学習手順

ADF は、ウェイトの付与と分岐条件の付与を全決定木で並列に指定の深さ  $D$  まで繰り返すことで決定木集合を生成する。いま、決定木の木数を  $T$ 、深さ  $d$  を  $d = 1, \dots, D$  とし、深さ  $d$  の決定木集合を  $\mathcal{M}_d$ 、その要素である各決

定木を  $m_{dt}$  と表す ( $1 \leq t \leq T$ )。まず、ウェイトの付与を行うため、決定木の学習データへの当てはまりの指標として純度を導入する。純度は各学習データに対して計算される値であり、ある学習データが到達した葉ノードにある、そのデータと同一カテゴリの学習データの割合と定義する。学習データ  $(\mathbf{x}_i, y_i)$  の決定木  $m_{dt}$  における純度を  $g_{m_{dt}}(\mathbf{x}_i, y_i)$  とすると、式 (1) で表すことができる。

$$g_{m_{dt}}(\mathbf{x}_i, y_i) = \frac{\sum_{j=1}^N \delta(y_i, y_j) \times \eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_j)}{\sum_{j=1}^N \eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_j)} \quad (1)$$

ただし、 $\delta(y_i, y_j)$  は  $y_i$  と  $y_j$  が一致した時に 1、そうでないときに 0 を返す指示関数、 $\eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_j)$  は  $\mathbf{x}_i$  と  $\mathbf{x}_j$  が決定木  $m_{dt}$  において同じ葉ノードに存在する場合に 1、そうでない場合に 0 を返す指示関数である。

ここで、全学習データの純度が高い場合に決定木集合の学習データへの当てはまりは良いと考えられる。そのため、ADF は各学習データの純度を用いて決定木集合の損失を導入し、この最小化を図る。 $\mathcal{M}_d$  の損失は  $L_d$  と定義し、この損失は指数損失関数や対数損失関数などを用いて与えることが可能である。ここで、指数損失関数を用いた場合の  $L_d$  は、以下のように表すことができる。

$$L_d = \prod_{i=1}^N \exp \left\{ - \sum_{t=1}^T g_{m_{dt}}(\mathbf{x}_i, y_i) \right\} \quad (2)$$

ADF はこの損失を最小化するために、勾配法 [5] を用いて損失を大きくしている学習データの純度を重点的に向上させることを深さごとに繰り返す。そのための手続きでは、損失の勾配をウェイトとして各学習データに付与し、分岐条件付与の度に各学習データの純度を向上させる度合いを調整する。深さ  $d$  における学習データ  $(\mathbf{x}_i, y_i)$  に付与されるウェイトを  $w_d(\mathbf{x}_i, y_i)$  とし、損失として式 (2) を用いるものとするれば、 $w_d(\mathbf{x}_i, y_i)$  は式 (3) で計算される。

$$w_d(\mathbf{x}_i, y_i) = \frac{\partial L_d}{\partial (\mathbf{x}_i, y_i)} = \exp \left\{ - \frac{1}{T} \sum_{t=1}^T g_{m_{dt}}(\mathbf{x}_i, y_i) \right\} \quad (3)$$

ADF はこのウェイトを用いて、同じ深さにある各葉ノードに対して分岐条件を付与する。ここで各葉ノードへの分岐条件の付与は、ノード内にある学習データのウェイト総和を各カテゴリに対して算出し、それらから求めたエントロピーやジニ係数が、分岐条件付与によって最も向上する方法で行われる。これにより、純度が低くウェイトの大きい学習データの純度が分岐条件付与により大きく改善することとなる。このとき、個々の決定木が異なる特徴を持つように、各ノード分割の際にランダムに一部の説明変数群を割り当てて、その中から最良のものを分岐条件に用いる。

### 3.2 予測手順

ADF における予測は、全ての決定木の出力をアンサンブルすることで行う。ここで個々の決定木の出力は 1 つの新規データに対してカテゴリ数分あり、その出力は新規データ  $\mathbf{x}_{N+1}$  が到達した葉ノードにおける学習データの各カテゴリへの所属度合を表している。

新規データのカテゴリ予測時は、各決定木  $m_{Dt}$  の出力として、新規データに対応する葉ノードに付与されたカテゴリ  $k$  の学習データの割合を算出する。この出力を  $f_{m_{Dt}}(\mathbf{x}_{N+1}, k)$  とすると、以下の式 (4) で算出される。

$$f_{m_{Dt}}(\mathbf{x}_{N+1}, k) = \frac{\sum_{i=1}^N \delta(y_i, k) \times \eta_{m_{Dt}}(\mathbf{x}_{N+1}, \mathbf{x}_i)}{\sum_{i=1}^N \eta_{m_{Dt}}(\mathbf{x}_{N+1}, \mathbf{x}_i)} \quad (4)$$

さらに、全ての決定木の出力をカテゴリ毎に平均化することで、新規データのカテゴリ  $k$  への所属度合を算出する。これを  $F_D(\mathbf{x}_{N+1}, k)$  とすると、その計算式は式 (5) のようになる。最終的に、新規データの予測カテゴリは式 (5) が最大となるカテゴリとする。

$$F_D(\mathbf{x}_{N+1}, k) = \frac{1}{T} \sum_{t=1}^T f_{m_{Dt}}(\mathbf{x}_{N+1}, k) \quad (5)$$

### 3.3 ADF の問題点

ADF には 2 つの問題点がある。1 点目は外れ値に対しても大きなウェイトを付与し、決定木集合が外れ値の影響を受けやすくなり、過学習を起こす可能性があることである。2 つ目は、分岐条件付与のたびに候補となる説明変数群を割り当て直すため、学習データへの当てはまりを過度に抑制し、交互作用の表現などが難しくなるなど、全体の予測精度が低下する可能性があることである。

## 4 提案手法

### 4.1 概要

本研究では、各学習データの外れ値である度合いの推定と活用および、原則的に決定木の生成で根ノードに割り当てた説明変数群を繰り返し用いることを許容し、必要に応じてその説明変数群の再割り当てを行う手法を提案する。また、外れ値の推定と説明変数の再割り当てを行う決定木の選択のため、各決定木に対して学習時に使用しない学習データ (OOB データ) を生成し、それらを活用する。OOB データは各決定木にとっては新規データと同じ性質を持つといえ、OOB データを生成途中の決定木で予測することで、各決定木の汎化性能や各学習データの外れ値である度合いを見積もることができる。提案手法では、見積もった外れ値の度合いや決定木の性能をウェイト算出や割り当てた説明変数の変更に活用する。すなわち、提案手法は以下の方法を組み合わせた手法である。

1. ブートストラップ法の導入
2. ウェイト付与方法の改善
3. 決定木への説明変数割り当て及び再割り当て

### 4.2 ブートストラップ法の導入

各決定木に対して学習に用いないデータセットを割り当てるために、データセットを学習用データセットと OOB データセットに 2 分割する。本研究では、そのための方法としてブートストラップ法 [6] を用いる。これは決定木のアンサンブル手法として一般的な Random Forests などを用いられている手法で、各決定木に対して重複を許してランダムに学習データを抽出することを学習データ数分繰り返し、サブデータセットを生成する。これを用いることで、学習データから学習に用いるサブデータセットを生成すると同時に、各決定木に対して学習に用いないデータ (OOB データ) を用意することが可能となる。また、各学習データに着目すると、そのデータを学習データとして用いている決定木と OOB データとして扱う決定木に分けることができる。

ここで、ADF における損失は、決定木集合の学習データへの当てはまりとして導入されているため、損失やウェイトを算出する際に、その学習データを OOB データとして扱う決定木における純度を考慮することは望ましくない。そのため、従来の損失やウェイトの算出方法を変更する必要がある。すなわち、ブートストラップサンプルを用いた場合の損失は学習に用いられた決定木のみにおける純度の平均から計算するべきであり、それを用いてウェイトを算出することが望ましい。ここで、学習時に用いられた決定木における純度の平均を  $g^{(w)}(\mathbf{x}_i, y_i)$  と

すると、深さ  $d$  における学習データ  $(\mathbf{x}_i, y_i)$  のウェイトは式 (6) で求めることができる。

$$w_d(\mathbf{x}_i, y_i) = \exp\{-g^{(u)}(\mathbf{x}_i, y_i)\} \\ = \exp\left\{-\frac{1}{|\mathcal{U}_i|} \times \sum_{t \in \mathcal{U}_i} g_{m_{dt}}(\mathbf{x}_i, y_i)\right\} \quad (6)$$

ただし、 $\mathcal{U}_i$  を学習データ  $(\mathbf{x}_i, y_i)$  を学習に用いた決定木集合とし、 $|\cdot|$  を集合に含まれる要素数とする。

### 4.3 ウェイト付与方法の改善

従来手法は、純度が低い学習データに対して大きいウェイトを付与することで、全ての学習データを均等に予測できる決定木集合の生成を目指す。しかし、外れ値に対し、大きなウェイトを付与してしまうことで過学習が生じることがある。そこで、本研究では外れ値のウェイトを相対的に小さくするようにウェイトに正規化項を加える。

ここで、各学習データに対して、そのデータを OOB データとして扱う決定木のみを用いてカテゴリ予測した際、この結果が悪い学習データは外れ値であると考えられる。そこで提案手法では、その結果から各学習データの外れ値の度合いを推定し、外れ値でない学習データのウェイトが大きくなるように、正規化項を加える。ここで、データ  $(\mathbf{x}_i, y_i)$  に対する、そのデータを学習時に用いない決定木における純度平均を  $g^{(o)}(\mathbf{x}_i, y_i)$  とし、 $g^{(o)}(\mathbf{x}_i, y_i)$  を式 (7) により算出する。

$$g^{(o)}(\mathbf{x}_i, y_i) = \frac{\sum_{i' \in \mathcal{O}_i} \delta(y_i, y_{j'}) \times \eta_{m_{dt}}(\mathbf{x}_{i'}, \mathbf{x}_i)}{\sum_{i' \in \mathcal{O}_i} \eta_{m_{dt}}(\mathbf{x}_{i'}, \mathbf{x}_i)} \quad (7)$$

ただし、 $\mathcal{O}_i$  は学習データ  $(\mathbf{x}_i, y_i)$  を学習に用いなかった決定木集合である。

ここで、各学習データのウェイトに付与する正規化項を  $g^{(o)}(\mathbf{x}_i, y_i)$  と損失関数を用いて求める。また、決定木間の相関の維持および効率的な損失の低減と外れ値の考慮をバランスさせるために、チューニングパラメータを導入し、純度から求めるウェイトと正規化項を組み合わせる比率を決定木ごとに変化させる。ここで、提案手法において学習データ  $(\mathbf{x}_i, y_i)$  に付与する決定木  $m_{dt}$  でのウェイトを  $W_{dt}(\mathbf{x}_i, y_i)$  とする。正規化項に指数損失関数を用いた場合、 $W_{dt}(\mathbf{x}_i, y_i)$  を式 (8) のように計算する。

$$W_{dt}(\mathbf{x}_i, y_i) = \exp\{-g^{(u)}(\mathbf{x}_i, y_i)\} \\ + C_t \cdot \exp\left\{-\{1 - g^{(o)}(\mathbf{x}_i, y_i)\}\right\} \quad (8)$$

ただし、 $C_t$  はチューニングパラメータとする。

提案手法では、決定木  $m_{dt}$  における学習に用いたデータの純度平均を  $PUR_t$  とし、OOB データの純度平均を  $OOBE_t$  とした上で、これらを用いてチューニングパラメータ  $C_t$  を決定する。ここで、 $PUR_t$  は式 (9) で計算し、 $OOBE_t$  は式 (10) で計算する。

$$PUR_t = \frac{1}{|\mathcal{U}_i|} \times \sum_{i' \in \mathcal{U}_i} \frac{\sum_{i \in \mathcal{U}_i} \delta(y_i, y_{j'}) \times \eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_{i'})}{\sum_{i \in \mathcal{U}_i} \eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_{i'})} \quad (9)$$

$$OOBE_t = \frac{1}{|\mathcal{O}_i|} \times \sum_{i' \in \mathcal{O}_i} \frac{\sum_{i \in \mathcal{U}_i} \delta(y_i, y_{j'}) \times \eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_{i'})}{\sum_{i \in \mathcal{U}_i} \eta_{m_{dt}}(\mathbf{x}_i, \mathbf{x}_{i'})} \quad (10)$$

これらを用いて、チューニングパラメータ  $C_t$  を以下のように設定する。

$$C_t = \frac{PUR_t}{OOBE_t} \quad (11)$$

このようにチューニングパラメータ  $C_t$  を決定することで、未だ汎化性能が低く、外れ値ではない学習データを重視した方が良い決定木においては正規化項を大きくし、すでに汎化性能が高い決定木においては過度に正規化項を大きくしすぎないようにする。これにより、過学習を防ぎつつ全体の損失を効率的に減少させることと、決定木間の相関を小さくすることが可能となる。

### 4.4 決定木への説明変数割り当て及び再割り当て

従来手法では、分岐条件を付与する際に候補とする説明変数をノードごとに制限することで、学習データへの当てはまりを抑制し、決定木間の相関を下げている。しかし、学習データへの当てはまりを過度に抑制し、全体の予測精度が下がることが考えられる。そこで、提案手法は学習データへの当てはまりを過度に抑制することを避けるため、必要が生じるまで一度割り当てた説明変数群のみを各木の生成に用いる。また、同一の説明変数群を用い続けることによる過学習や表現能力低下による予測精度の低下を防ぐため、以下の手順で学習途中で一部の決定木に対して割り当てた説明変数を変更する。

1. 割り当てた説明変数を変更する決定木を決定する。
2. その決定木に割り当て直す説明変数を選択する。

提案手法では、学習途中で割り当てた説明変数に問題がある決定木を発見するために OOB データの誤り率を用いる。一般に、OOB データに対する予測性能はその決定木の汎化性能と解釈でき、汎化性能の低い決定木は割り当てた説明変数を用いて、有効な分岐条件を付与することが困難であると考えられる。そこで深さ毎に OOB データの誤り率が高い  $\alpha\%$  の決定木に対し、割り当てられた説明変数群の再割り当てを行う。この再割り当てにより、予測精度が低い決定木に対して予測に寄与する分岐条件が付与され、予測精度の向上を図ることが可能となる。

また、提案手法では説明変数を割り当て直す際、各木の相関をなるべく上げずに交互作用効果の影響を考慮できるような説明変数の組み合わせを選択する。ここでは、再割り当ての段階において各決定木に割り当てられている説明変数の組み合わせを考慮する。すなわち、説明変数の再割り当てを行う決定木にすでに割り当てられている説明変数となるべく同時に用いられていない説明変数を割り当てることにより、決定木間の相関をなるべく上げずに、説明変数間の様々な交互作用効果の考慮を行う。

### 4.5 提案手法アルゴリズム

ADF を用いた提案手法は学習手順と予測手順で構成されている。以下では両者のアルゴリズムについて述べる。

#### 【学習手順】

- Step1) 全ての学習データに対して均等にウェイトを付与し、各木に一部の説明変数を割り当てた上で、深さ  $d = 1$  の決定木を  $T$  本生成する。
- Step2) 全ての決定木における、深さ  $d$  のノードに分岐条件を付与し、深さ  $d + 1$  の決定木集合を生成。
- Step3) 式 (8) を用いて学習データのウェイトを更新。
- Step4) OOB データの平均純度が低い  $\alpha\%$  に対し、割り当てた説明変数群を変更する。
- Step5) 予め指定した深さ  $D$  に到達したら終了。そうでなければ Step2 へ戻る。

#### 【予測手順】

- Step1) 各決定木を用いて新規データ  $\mathbf{x}_{N+1}$  を葉ノードまで到達させ、式 (4) から出力を計算。
- Step2) 各決定木の出力を平均化し、ADF の出力を計算。
- Step3) 式 (5) から新規データ  $\mathbf{x}_{N+1}$  のカテゴリを予測。

□

## 5 実験

### 5.1 UCI データセットを用いた評価実験

提案手法の有効性を示すため、UCI 機械学習レポジトリ [4] のデータを用いた ADF との予測精度の比較を行う。

#### 5.1.1 実験条件

実験では式 (12) で示される誤り率を用いて評価した。

$$\text{誤り率} = 1 - \frac{\text{正しく分類されたテストデータ数}}{\text{テストデータ数}} \quad (12)$$

実験では各データセットに対して、ランダムに学習データとテストデータを入れ替えて 50 回実験を行い、その平均を計算した。実験に用いるパラメータは予備実験から、 $T = 200$ ,  $D = 15$ ,  $\alpha$  は交差検定でデータセットごとに [30, 40, 50, 60, 70] の中から選択することとした。また、実験に用いたデータセットの概要を次の表 1 に示す。

表 1. データセットの概要

データセット	次元数	カテゴリ	学習データ	テストデータ
index	13	2	550	200
CNAE-9	856	9	580	500
LSVT	310	2	80	46
ionosphere	34	2	200	150
sensor	24	4	1457	4000

#### 5.1.2 実験結果と考察

図 2 に実験の結果を示す。その際、従来手法と提案手法で統計的有意差があるか  $t$  検定を行う。図 2 における\*\*は 1% 有意、\*は 5% 有意を示している。

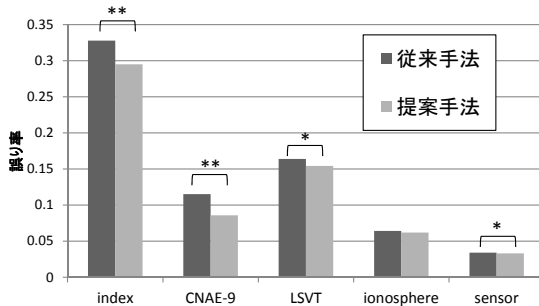


図 2. 実験結果

全ての実験において、従来手法と比較し、提案手法が同等以上の精度を示している。また、データセット index, CNAE-9, LSVT と sensor においては有意差がある結果となった。しかし、データセット ionosphere において有意差はなく、sensor に関しても提案手法の効果小さい。このことから、提案手法は高次元なデータセットや従来手法で精度が低いデータセットに対して有効であると考えられる。これらのデータセットは説明変数間の相互作用が多く存在すること、モデルでの表現が難しいことや外れ値の割合が大きいため、提案手法における学習データへの当てはまり向上や外れ値を考慮したウェイト付与の効果が大きかったと考えられる。

### 5.2 外れ値に対する頑健性に関する実験

本節では提案手法の外れ値に対する頑健性を評価するため、人工的に外れ値を付加したデータを用いて検証する。

#### 5.2.1 実験条件

5.1 の実験で用いたデータセット中で予測精度が高く外れ値が少ないと考えられる sensor を用いて、ランダムに学習データのカテゴリを変更することにより、人工的に外れ値を生成し予測精度への影響を検証する。ここでは

提案手法と従来手法の誤り率を比較する。実験は、カテゴリを変更する学習データの数を 10 個ずつ増加させ、予測精度の推移を検証する。また、各実験について 10 回実験を行い、平均値を計算した。

#### 5.2.2 実験結果と考察

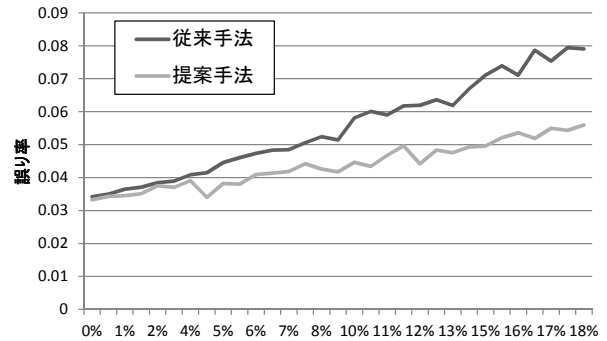


図 3. 外れ値と精度の関係

図 3 に外れ値と予測精度に関する実験の結果を示す。人工的に発生させた外れ値が少ないうちは従来手法と提案手法の精度に差は小さいことがわかる。しかし、外れ値の割合が増えると、従来手法は提案手法より急速に精度が悪化することがわかる。これは以下の 2 つの理由が考えられる。1 つ目は提案手法が外れ値の度合いを学習に導入したことによって外れ値の影響を小さくし、過学習を防止したことである。2 つ目は提案手法が説明変数を各木で原則的に固定したため、決定木間の相関が下がり汎化性能が向上したと考えられる。以上により、提案手法が特に外れ値などに強い手法であるといえる。

## 6 まとめと今後の課題

本研究では、各決定木に対して学習に用いないデータを割り当て、決定木を成長させる度にこれらのデータの予測精度を図り、この結果を利用してこれら ADF の予測精度を改善した。また、実験の結果から提案手法が従来手法の ADF よりも外れ値に強いことと、モデル表現が難しいデータや説明変数の数が多いデータに対して優れていることが分かり、提案手法の有効性が示された。今後の課題は、変数再選択法に用いる最適なパラメータ  $\alpha$  の自動決定や提案手法に対する解析的な検証が挙げられる。

## 参考文献

- [1] L. Breiman, J. Freidman, R. Olshen, and C. Stone, "Classification and Regression Trees," *IWadsworth*, Belmont, CA, 1984.
- [2] L. Breiman, "Random Forests," *Machine Learning*, Vol.45, pp.5-32, 2001.
- [3] S. Schuster, P. Wohlhrst, C. Leistner, A. Saffari, P. Roth and H. Bischof, "Alternating Decision Forests," *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp.508-515, 2013.
- [4] C. Blake, C. Merz, "UCI repository of machine learning database.," <http://www.ics.uci.edu/mllearn/mlrepository.html>, 1998.
- [5] 杉山将, "統計的機械学習," 株式会社オーム社, 2009.
- [6] 金森敬文, 畑木晃平, 渡辺治, "ブースティング-学習アルゴリズムの設計技法-, 森北出版株式会社, pp.26-50, 2006.