

アンサンブル学習 pdi-BoostingG の定式化とその評価

A Proposal of Ensemble Learning “pdi-BoostingG” and its Evaluation

○ 入江 穂乃香, 林 勲
○ Honoka Irie, Isao Hayashi
関西大学大学院
Kansai University

Abstract: Recently, we have been increasing interest in ensemble learning. In particular, Boosting is a useful learning method in which multiple classifiers construct multiple layers with maintaining mutual dependencies, and its expressive ability is high. We have proposed pdi-Boosting, which virtually generates interpolated data and adds them to the training data to improve the accuracy of classifiers. However, in this method, the area for generating the additional fuzzy rule is not directly related to the misidentification data and virtual data. Therefore, the discrimination rate is not high. In this paper, we propose a new “pdi-BoostingG” that also regularizes the additional region G of fuzzy rules. In pdi-BoostingG, fuzzy rules are additionally generated around the misidentification data, and virtual data is generated in this area G. In addition, the individual learning of the additional fuzzy rule and the whole learning of the whole space are alternately learned to form the multi-layered structure. As a result, pdi-BoostingG improves the discrimination rate and robustness. In addition, since virtual data and fuzzy rules are inherited between multiple layers, deep inference of fuzzy rules is realized. We formulate here how to generate additional fuzzy rules and virtual data, and how to inherit whole fuzzy rules and virtual data between layers. In addition, we compare the discrimination rate of pdi-BoostingG with other methods in order to discuss its usefulness.

1. はじめに

パターン分類問題に対して、複数個の弱判別器を統合的に組み合わせる全体の識別精度を向上させるアンサンブル法 [1,2] に対する関心が高まっている。アンサンブル学習の判別器結合モデルで逐次的タイプの一つにブースティング法 (Boosting Methods) [3] がある。ブースティングは、逐次的に弱判別器を学習させ、識別率を向上させる手法であり、Adaboost [3,4] は特に有用で、データ集合の特徴量を解析しやすい利点がある。また、学習データの分割法 [5] の提案や 2 クラス識別器の Gentle Boosting [6] をマルチクラスに拡張した Joint Boosting [7] なども提案されている。なお、バギングとブースティングを統合したアンサンブル法も提案されている [8,9]。我々は弱判別器にファジィ推論を用いて、そのモデルの学習の際に、仮想的に生成したデータを観測データに追加して学習データを構成し、判別線の精度を向上させる新たなバギング法やブースティング法を提案している。このバギング法を pdi-Bagging と呼び、仮想的に生成したデータをバーチャルデータと呼んでいる [10–14]。バーチャルデータの追加で学習時にデータ量が増えるので、クラス間でデータ量の偏りがなくなり、判別線の同定精度が向上する。しかし、バギングでは、最適化したファジィルールや発生したバーチャルデータが次層に引き継がれず、全体の認識精

度が改善されない。そこで、多層構造における前層から次層にファジィルールやバーチャルデータ集合を継承するブースティング法を提案している [15]。このブースティング法を pdi-Boosting (Possibilistic Data Interpolation-Boosting) と呼ぶ。ブースティングでは、前層で学習された判別器のファジィルールを次層に引き継ぐルール継承を行う。また、バーチャルデータを次層の学習データに追加して学習データの量を増加させるバーチャルデータの量的継承を行う。さらに、前層から次層へ継承するバーチャルデータクラスを再決定する方法やサンプリングの方法などのバーチャルデータの質的継承も行う。しかし、ブースティング法では、バーチャルデータを発生する領域に追加でファジィルールを設定するが、その設定位置は任意で決定するため、最終的な識別率は必ずしも良いとはいえない。

本論文では、従来型の pdi-Boosting に加えて、ファジィルールの追加領域の位置を規則化して決定する新たな pdi-Boosting を提案する。この手法を pdi-BoostingG と呼ぶ。pdi-BoostingG では、誤判別データが識別された際、その誤判別データを中心に範囲 G を設定して、範囲 G 内にファジィルールを追加で生成し、バーチャルデータを範囲 G 内で発生する。全体のアルゴリズムでは、範囲 G 内の追加ファジィルールの個別学習と全空間のファジィルール

の全体学習を交互に繰り返して多層構造を形成する。まず、全空間の全体学習では、層構造において前層から次層へ移行する際にファジールール集合とバーチャルデータ集合を継承する。ファジールールの継承では、全空間のファジィ推論の学習において、前層の学習後のファジールールは次層のファジールールの学習の初期値として用いられる。また、バーチャルデータの継承では、前層で発生したバーチャルデータは、継承率を乗じた量が次層でバーチャルデータとして用いられる。次に、範囲 G の個別学習では、誤判別が識別された際、その誤判別データを中心に範囲 G を設定し、範囲内でバーチャルデータを発生して、範囲内のファジールールを学習する。なお、範囲 G では、直角台形型のメンバシップ関数を用いるので、範囲 G の大きさは変化せず、誤判別に特化したファジールールが追加される。ただし、層構造の前層から次層へ移行する際には、範囲 G でファジールールは継承するが、バーチャルデータは全空間のファジィ推論の学習の際にすでに継承されているので、範囲 G では継承しない。

このように、全空間と範囲 G の 2 種類の学習機能と多層構造によるファジールールとバーチャルデータの継承により最終結果の識別率は向上する。なお、本手法は、範囲 G 内でバーチャルデータを発生して、全空間との多層構造でファジールールとバーチャルデータの両方を継承することから、一般的な多段ファジィ推論とは異なる。ここでは、範囲 G のファジールールとバーチャルデータの設定方法、及び、ファジールールとバーチャルデータの継承方法を定式化する。また、2 群識別の数値例を用いて、pdi-BoostingG の特性を明らかにして、他手法との精度比較を行い、本手法の有用性を議論する。

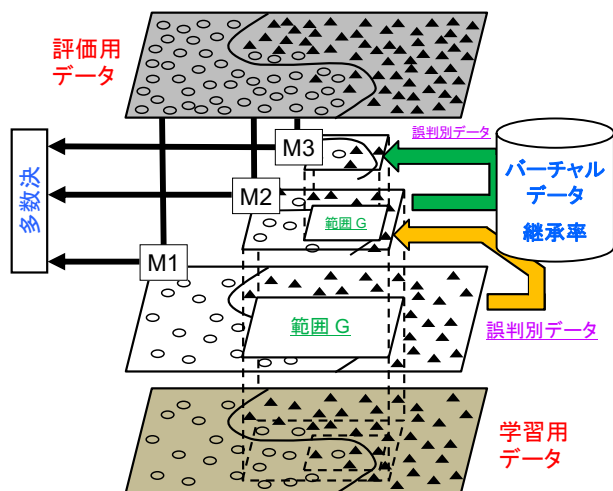


図 1: pdi-BoostingG

2. pdi-BoostingG

pdi-BoostingG の概念を図 1 に示す。pdi-BoostingG では、全データ集合から確率的に抽出された学習データ (TRD) を用いてファジィ推論の弱判別器 M_0 を学習し、 TRD の識別率を算出する。次ステップ (層) では、誤判別データを中心とした範囲 G を設定し、メンバシップ関数を用いてバーチャルデータを範囲 G 内で発生して継承率を乗じた量を TRD に追加させて、ファジィ推論の弱判別器 M_1 により TRD の識別率を算出する。 TRD を増加させることによって弱判別器の識別精度が向上する。終了判定が満足されるまでこの一連の操作を L 回繰り返し、最終的に、評価データ (CHD) を L 個の弱判別器 M_0, M_1, \dots, M_L に入力して、多数決により最終結果を得る。

pdi-BoostingG では、弱判別器として簡易型ファジィ推論 [16, 17] によるファジィクラスタリングを用いる。ただし、三角型のメンバシップ関数を一般化した正規な台形型ファジィ集合を用いる。いま、出力変数を z 、後件部のシングルTONを p_i で表すと、ファジールール $r_i, i = 1, 2, \dots, R$ は次のようになる。

$$r_i : \text{if } x_1 \text{ is } \mu_{F_{i1}}(x_1) \text{ and } \dots \text{ and } x_n \text{ is } \mu_{F_{in}}(x_n) \\ \text{then } C = \{C_{ik} \mid z = p_i\}$$

ただし、 C は出力クラスの変数であり、 C_{ik} はルール r_i のクラス値が C_k であることを示す。

いま、入力データ $\mathbf{x} = (x_1, x_2, \dots, x_n)$ が得られたとしよう。第 i 番目のファジールール r_i の前件部に入力データ \mathbf{x} を入力し、前件部の適合度 $\mu_i(\mathbf{x}) = \mu_{F_{i1}}(x_1) \cdot \mu_{F_{i2}}(x_2) \cdot \dots \cdot \mu_{F_{in}}(x_n)$ を計算する。ファジィ推論の結果 \hat{z} とクラス C は次式から求める。

$$\hat{z} = \frac{\sum_{i=1}^R \mu_i(\mathbf{x}) \cdot p_i}{\sum_{i=1}^R \mu_i(\mathbf{x})} \\ C = \{C_k \mid \min |\hat{z} - z|\}$$

W 個のデータからなるデータ集合 D の第 d 番目のデータを $\mathbf{x}^D(d) = (x_1^D(d), x_2^D(d), \dots, x_j^D(d), \dots, x_n^D(d))$ で表す。バーチャルデータの発生では、任意の学習データ $\mathbf{x}^D(d)$ のクラスが弱判別器により誤判別と判別された場合、その誤判別データ $\mathbf{x}^E(d)$ を中心に範囲 G を生成して、バーチャルデータ $\mathbf{x}^V(d)$ を発生する。 $\mathbf{x}^V(d)$ を構成する第 j 属性目のバーチャルデータ $x_j^V(d)$ の発生は、ある実数 $h, 0 \leq h \leq 1$ が与えられると、ファジィ数 F のメンバ

シップ関数 $\mu_F(x_j)$ を用いて次のように発生する.

$$x_j^V(d) = \{x_j \mid \mu_F(x_j) = h, \mu_F(x_j^E(d)) = 1\}$$

$$h \sim N(1, 1), \quad 0 \leq h \leq 1$$

ただし, $x_j^E(d)$ は誤判別データを意味し, メンバシップ関数 $\mu_F(x_j)$ は, ファジィ数 F の中心が $x_j^E(d)$ であり, 標準偏差が σ である次の正規分布で定義する.

$$\mu_F(x_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_j - x_j^E(d))^2}{2\sigma^2}\right)$$

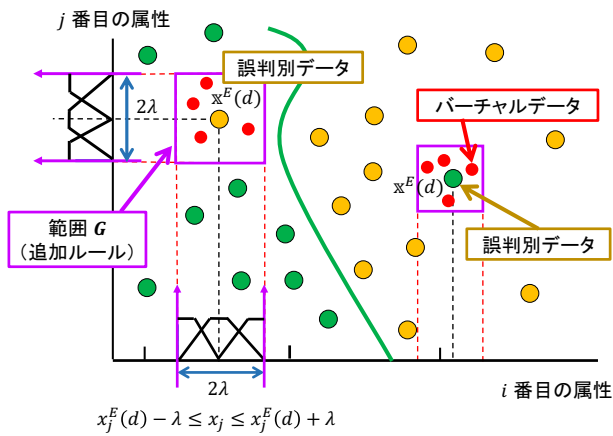


図 2: Generation of Virtual Data and Fuzzy Rules in Region G

3. 継承と範囲 G

識別率をより向上させるため, pdi-BoostingG において, 前層から次層にルールやデータ集合を継承する. ここでの継承問題とは, 次の 3 種類の継承を対象とする.

(1) 全空間ファジィルールの継承

判別器によって学習された前層の全空間のファジィルールを次層に引き継ぐための継承法をいう. 前層で識別率が最大となったファジィルールのメンバシップ関数のパラメータを次層のファジィルール学習の際の初期設定のパラメータとする. 前層の最適ファジィルールが次層で継続して学習されるので, ファジィルールがより深化に学習され識別率の向上が期待できる.

(2) 全空間のバーチャルデータの量的継承

各層で発生した全空間のバーチャルデータを次層の学習データに追加して学習データの量を増加させるための継承法をいう. 全空間のバーチャルデータの量的継承には 2 つの方法がある.

(a) バーチャルデータ量の単率継承法

第 l 層の v_l 個のバーチャルデータから継承率 ξ を用いた乱数サンプリングによって, v_{l+1} 個のバーチャルデータを第 $l+1$ 層へ継承する.

$$v_{l+1} = \xi \times v_l$$

$$W_{l+1}^{TRD} = W_l^{TRD} + v_{l+1}$$

ここで, W_l^{TRD} と W_{l+1}^{TRD} はそれぞれ第 l 層目と第 $l+1$ 層目の学習データである.

(b) バーチャルデータ量の複率継承法

第 1 層から第 l 層までの $\sum_{k=1}^l v_k$ 個のバーチャルデータから継承率 ξ を用いた乱数サンプリングによって, v_{l+1} 個のバーチャルデータを第 $l+1$ 層へ継承する.

$$v_{l+1} = \xi \times \sum_{k=1}^l v_k$$

$$W_{l+1}^{TRD} = W_l^{TRD} + v_{l+1}$$

(3) バーチャルデータの質的継承

層間で継承したバーチャルデータのクラスを再決定することとサンプリング方法をいう. 特に, バーチャルデータのクラスは前層の学習データやバーチャルデータを考慮して決定すべきである.

従来の pdi-Boosting では, バーチャルデータを誤判別データ周辺で発生しファジィルールは設定していない. そのため, 全空間のファジィルールは個々の誤判別データに適応できないと考えられる. そこで, pdi-BoostingG では, 誤判別データを中心に範囲 G を限定して生成し, この範囲 G 内でファジィルールを設定して, 全空間のファジィルールとは独立に学習させる. 図 2 に範囲 G におけるバーチャルデータとファジィルールの設定の例を表す. $x^D(d)$ が誤判別データ $x^E(d)$ と識別された場合, その $x^E(d)$ を中心に幅 2λ の範囲 $G \subset [0, 1]$ を生成し, その範囲 G 内でバーチャルデータ $x^{V(E)}(d)$ を発生する.

$$\mathbf{x}^{V(E)}(d) \in G$$

$$G = \{x_i \mid x_i^E(d) - \lambda \leq x_i \leq x_i^E(d) + \lambda\}, \quad 0 \leq \lambda \leq 1$$

ここで, λ は範囲 G の大きさを指定する定数である.

バーチャルデータ $\mathbf{x}^{V(E)}(d)$ は、範囲 G 内で $\mathbf{x}^E(d)$ を中心とする正規分布を用いた正規乱数で生成される。ただし、範囲 G を固定するため、範囲 G の両端では、直角台形型メンバシップ関数を設定し、範囲 G 内に設定した全てのメンバシップ関数は、学習後も範囲内に留まるように制限した。なお、 $\mathbf{x}^{V(E)}(d)$ は次層以降の範囲 G の学習では使用しないので学習後に破棄する。

pdi-Boosting における継承を図 3 に示す。pdi-Boosting を 3 層構造としている。各層で全空間のファジィルールと範囲 G 内の追加ファジィルールを用いたファジィ推論により判別線が推定されている。層間でファジィルールとバーチャルデータが継承されている。なお、継承率 ξ が $\xi = 0.0$ の場合には、pdi-Boosting は pdi-Bagging と同じ構造となることに注目すべきである。

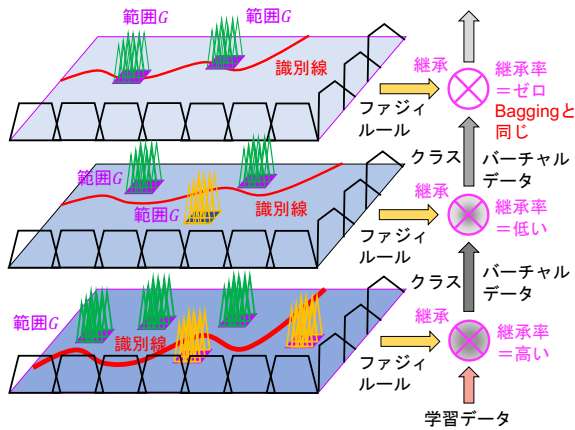


図 3: Inheritance Mechanism of pdi-BoostingG

pdi-Bagging のアルゴリズムを次のように定式化する。

- Step 1** 計測データ D (個数: W 個) を学習データ D^{TRD} (個数: W^{TRD} 個) と評価データ D^{CHD} (個数: W^{CHD} 個) に分割する。また、全空間を学習するための弱判別器のファジィルールの集合を Q で表し、 D^{TRD} から構成する全空間を学習するためのバーチャルデータを $D^V (= D^{V(C)} + D^{V(E)})$ で表す。ただし、 $D^{V(C)}$ は D^{TRD} の誤判別データ $\mathbf{x}^E(d)$ が弱判別器の学習により正判別となった場合のバーチャルデータを表し、 $D^{V(E)}$ は学習後も誤判別であった場合のバーチャルデータを表す。
- Step 2** 第 0 層目として、全空間のメンバシップ関数をもつ弱判別器 M_0 に D^{TRD} を入力し弱判別器 M_0 を学習する。第 0 層目の結果 R_0 の識別率 r_0^{TRD} を得る。学習された全空間のファジィルールを Q に追加する。

- Step 3** 第 l 層目での範囲 G の生成のため、 D^{TRD} の誤判別データ $\mathbf{x}^E(d)$ を抽出し、その $\mathbf{x}^E(d)$ が第 $l-1$ 層以前で抽出されたとしても、 $\mathbf{x}^E(d)$ を中心に範囲 $G \subset [0, 1]$ を生成する。範囲 G 内でバーチャルデータ $\mathbf{x}^{V(E)}(d)$ を生成し、 $\mathbf{x}^{V(E)}(d)$ を全空間の学習のための $D^{V(E)}$ に保存する。

$$\mathbf{x}^{V(E)}(d) \in G$$

$$G = \{x_l | x_l^E(d) - \lambda \leq x_l \leq x_l^E(d) + \lambda\}$$

$$0 \leq \lambda \leq 1$$

ただし、 λ は範囲 G の大きさを指定する定数である。

- Step 4** 範囲 G で弱判別器のファジィルールのメンバシップ関数を設定し、第 l 層目の弱判別器 M_l^G に $\mathbf{x}^E(d)$ と $\mathbf{x}^{V(E)}(d)$ 、範囲 G 内の D^{TRD} を入力し、弱判別器 M_l^G を学習する。ただし、範囲 G を固定するため、範囲 G の両端には、直角台形型メンバシップ関数を設定し、範囲 G 内に設定する全てのメンバシップ関数には、学習後も範囲内に留まる制限が設定される。 $\mathbf{x}^{V(E)}(d)$ は次層以降の範囲 G の学習では使用しないので破棄する。学習されたファジィルールを全空間の学習のための Q に追加する。

- Step 5** 第 l 層目での全空間の弱判別器の学習のため、弱判別器 M_l に D^{TRD} と $D^{V(C)}$ 、 $D^{V(E)}$ を入力し、 Q に存在する全てのファジィルールを用いて弱判別器 M_l を学習し、 Q の全ルールを更新する。

- Step 6** 第 l 層目の全空間の弱判別器の学習によって、正判別データ $\mathbf{x}^C(d)$ が誤判別となった場合、過去に $\mathbf{x}^C(d)$ から生成した $\mathbf{x}^{V(C)}(d)$ も含めて全ての $\mathbf{x}^{V(C)}(d)$ を $D^{V(C)}$ から $D^{V(E)}$ へ移動する。一方、誤判別データ $\mathbf{x}^E(d)$ が正判別となった場合、過去に $\mathbf{x}^E(d)$ から生成した $\mathbf{x}^{V(E)}(d)$ も含めて全ての $\mathbf{x}^{V(E)}(d)$ を $D^{V(E)}$ から $D^{V(C)}$ へ移動する。なお、過去の $\mathbf{x}^E(d)$ が再び誤判別 $\mathbf{x}^E(d)$ となり、かつ、単率継承の場合には、 $\mathbf{x}^{V(E)}$ は $D^{V(E)}$ から取り除くが、複率継承では、 $\mathbf{x}^{V(E)}$ は $D^{V(E)}$ に存続させる。ただし、 $\mathbf{x}^{V(E)}$ は発生した第 l 層目から第 $l + \phi$ 層目まで存続するとし、第 $l + \phi + 1$ 層目で $D^{V(C)}$ と $D^{V(E)}$ から取り除く。 D^{TRD} に対する第 l 層目の結果 R_l の識別率 r_l^{TRD} を得る。

Step 7 第 l 層目のバーチャルデータ $D^{V(C)}$ と $D^{V(E)}$ に対して, 全ての $\mathbf{x}^{V(C)}$ と $\mathbf{x}^{V(E)}$ のクラス k^* を求める.

$$k^* = \{k | \min_k E^k = \min_k (w_1 E_1^k + w_2 E_2^k + w_3 E_3^k)\}$$

ただし, E_1^k, E_2^k, E_3^k は, それぞれクラス決定のための評価基準であり, E^k は全体の評価値である. また, w_1, w_2, w_3 は各評価値の重みである.

Step 8 第 l 層目での v_l^C 個の $D^{V(C)}$ と v_l^E 個の $D^{V(E)}$ が, 継承率 ξ^C と ξ^E と乱数サンプリングによって, 第 $l+1$ 層目では, それぞれ, v_{l+1}^C 個の $D^{V(C)}$ と v_{l+1}^E 個の $D^{V(E)}$ に継承される.

$$v_{l+1}^C = \xi^C \times v_l^C$$

$$v_{l+1}^E = \xi^E \times v_l^E$$

Step 9 $l = l+1$ として, Step3 から 8 までを繰り返し, しきい値 θ に対して $r_l^{CHD} \geq \theta$ を満足した $K = l$ の時点, あるいは, 弱判別器の個数 L と繰り返し回数 K , $K \leq L$ に対して $l \geq K$ を満足した時点でアルゴリズムを終了する.

Step 10 M_K に D^{CHD} を適用し, 多数決により最終の結果の識別率 r_K^{CHD} を得る.

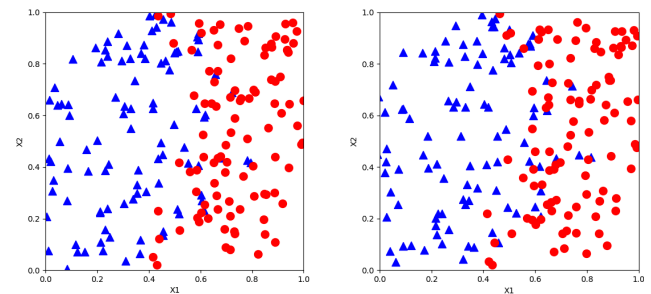
4. 数値データによる検証と考察

検証に用いる数値データは, 学習データと評価データでそれぞれ 200 個である. 学習データと評価データに用いる数値データを図 4 に示す. これらの数値データは, 基本データに乱数で最大 ± 0.05 の増減値を加えて生成した. ここでは, 2 入力 1 出力の 2 群判別問題として, ファジィ推論の 2 群クラスの実数値を 2.0 (赤・○印) と 3.0 (青・△印) に設定した. 弱判別器には簡易型ファジィ推論を用いる. 各入力区間 $[0, 1]$ に 5 種類の台形型メンバシップ関数を設定するので, 全空間では 25 個のルールとなる. 範囲 G では, 3 種類のメンバシップ関数を台形型で設定し, 範囲 G 内の各次元の両端では直角台形型メンバシップ関数を設定し, 範囲 G 内のメンバシップ関数は学習後も範囲内に留まる制限を設ける. この結果, 範囲 G の 9 個のルールは領域内で集中的に学習される.

ファジィ推論の前件部の初期値設定は既定法とし, 前件部と後件部の学習順序は, 後件部 \rightarrow 前後件部交互学習とする. 学習では, 各入力の台形型メンバシップ関数の上底の 2 頂点の x 座標 x_b と x_c , および, 上底と下底の x 座標の差 α と β の学習係数 $K_b, K_\alpha, K_c, K_\beta$ [17] を

同一とし 0.01 に設定する. また, 後件部のシングルトンの学習係数 K_p は, 最初の後件部学習では 0.4 に設定し, 交互学習の後件部学習では 0.6 とする. 後件部と前後件部交互学習のエポック回数をそれぞれ 10 回と (10, 10) 回に設定した.

バーチャルデータ発生時のメンバシップ関数 $\mu_F(x_j)$ は正規分布とし, バーチャルデータの発生個数は 1 個とする. バーチャルデータのクラス推定のための評価値の重みを $(w_1, w_2, w_3) = (0.5, 0.25, 0.25)$ とする. アルゴリズムの終了規範は繰り返し判定として, 回数は $K = 5$ とする. なお, 継承率は $\xi = 0.6$ とした, また, 結果の識別率は 10 試行の識別率の平均値とした.



(a) 学習データ

(b) 評価データ

図 4: Training Data and Testing Data

pdi-BoostingG の範囲 G のレベル λ を変化させた場合の識別率と他手法との比較を表 1 に示す. 表 1 では, 誤判別データが次層でも誤判別と識別された場合, そのルールの設定と追加方法, 及び, バーチャルデータの再発生の方法を 5 種類で変更して比較した. 具体的には, 範囲 G 内の誤判別データが次層でも再度誤判別と識別された場合, 範囲 G 内のファジィルールを初期化して設定するか, あるいは, 前層のルールを継承するか, また, その場合に, ファジィルールを新たに追加するか, あるいは, 非追加とするか, さらに, その範囲 G 内でバーチャルデータ (VD) を再度発生するか, あるいは, 非再発生とするかの方法を 5 種類に分類して識別率を求めた. 本論文の提案アルゴリズムは下線で示した (ルール継承, ルール非追加, VD 再発生) である. また, 従来の pdi-Boosting と pdi-Bagging, Adaboost とも比較した. なお, 判別器のファジィ推論単体の識別率は 89.00% であった. 一方, 範囲 G の大きさを設定する定数 λ を 0.01~0.1 で変化させた場合の識別率を図 5 に示す. 表 1 と図 5 の結果から, 識別率には次の特性があることがわかる.

表 1: Discriminant Rates Comparison between pdi-BoostingG and Other Methods

範囲 G λ	pdi-BoostingG					pdi-Boosting	pdi-Bagging	Adaboost J48	Adaboost RandomForest
	ルール初期化 ルール非追加 VD 再発生	ルール継承 ルール非追加 VD 非再発生	ルール継承 ルール非追加 VD 再発生	ルール継承 ルール追加 VD 非再発生	ルール継承 ルール追加 VD 再発生				
0.025	90.90	90.85	91.45	91.55	90.95	89.25	90.55	83.00	83.00
0.05	91.65	92.40	92.60	92.40	92.20	89.25	90.55	83.00	83.00
0.075	90.80	90.45	90.40	90.80	90.75	89.25	90.55	83.00	83.00
0.1	90.69	90.60	91.30	90.55	90.60	89.25	90.55	83.00	83.00
平均	91.00	91.08	91.44	91.33	91.13	89.25	90.55	83.00	83.00

- 1) pdi-BoostingG や pdi-Bagging, pdi-Boosting のバーチャルデータを発生する手法と Adaboost を比較すると、バーチャルデータを発生する手法の方が5%以上も識別率が高い。この結果から、バーチャルデータの発生の効果が現れているといえる。
 - 2) pdi-Bagging や pdi-Boosting と pdi-BoostingG の平均認識率を比較すると、pdi-BoostingG の方が識別率が高い。この結果から、範囲 G でのルールの発生とバーチャルデータの発生が識別率の向上に有用であるといえる。全空間のファジィ推論では、全ての学習データを用いて学習するため、特異点の学習データが無視される傾向がある。一方、範囲 G の生成によって、誤判別データ周辺は局所的に特異点となるクラスを構成することが可能となった。本手法は、階層構造でこれらの2種類の範囲 G と全空間の学習が交互に繰り返されるので、個々の誤判別データが減少することで識別精度が向上したといえる。
 - 3) pdi-BoostingG の中では、範囲 G 内の誤判別データが次層でも再度誤判別と識別された場合、範囲 G 内のファジィルールを初期化して発生する場合の識別率は $\lambda = 0.05$ の際に 91.65%と最大となり、前層のルールを継承する場合は、 $\lambda = 0.05$ で 92.20%以上となっている。したがって、継承する方式の方が識別率が高いといえる。これは、範囲 G 内での継続的なルール学習が全体の識別率を向上させていると考えられる。
 - 4) pdi-BoostingG の中では、範囲 G 内の誤判別データが次層でも再度誤判別と識別された場合、範囲 G 内でファジィルールを継承追加して発生させる場合には、平均識別率は 91.13%~91.33%となったが、ファジィルールを追加しない場合には、識別率は 91.08%~91.44%となり、若干、ルールを追加しない方式の方が識別率が高い。
 - 5) 「ルール非追加」でのバーチャルデータの再発生と非再発生の識別率を比較すると、ほぼバーチャルデータ再発生の方が識別率が高い。一方、「バーチャルデータ非再発生」でのルール追加とルール非追加を比較すると、その識別率には大きな差はない。一方、「ルール非追加+バーチャルデータ再発生」と「ルール追加+バーチャルデータ再発生」を比較すると、ルール非追加の方が識別率が高い。したがって、ルールは追加せず、バーチャルデータのみを再発生させる方式が良いと考えられる。この方式は本論文での提案アルゴリズムである。
 - 6) pdi-BoostingG の中で、最も高い識別率は $\lambda = 0.05$ の(ルール継承, ルール非追加, バーチャルデータ再発生)の 92.60%であり、この組み合わせは、本論文の提案アルゴリズムである。
 - 7) 範囲 G レベルの λ の変化に対する識別率は変動している。したがって、識別率は範囲 G の大きさに依存していることがわかる。
- これらの考察から、pdi-BoostingG は他手法よりも有用であり、範囲 G の生成が識別率に大きく影響を与えていることがわかる。また、pdi-BoostingG では、範囲 G 内にファジィルールは追加せず前層から継承し、バーチャルデータのみを発生させる方式が最も有用な手法であることがわかった。一方で、識別率は、範囲 G の大きさを示す λ の変化に対して変動する。したがって、ここでは、範囲 G レベル λ の変化に対する提案手法の識別率を議論する。

図 5 に、範囲 G レベル λ の $0.01 \leq \lambda \leq 0.1$ の変化に対する提案手法の識別率を示す。ここでは、提案アルゴリズムの識別率を青線で示し、比較のため、(ルール初期化・ルール非追加・VD 再発生) の識別率を赤線で示した。(ルール初期化・ルール非追加・VD 再発生) の識別率が、 $\lambda = 0.04$ でピークを示し最大識別率 92.00% を得たのに対して、提案アルゴリズムの (ルール継承・ルール非追加・VD 再発生) では、 $\lambda = 0.05$ でピークを示し最大識別率 92.60% を得た。

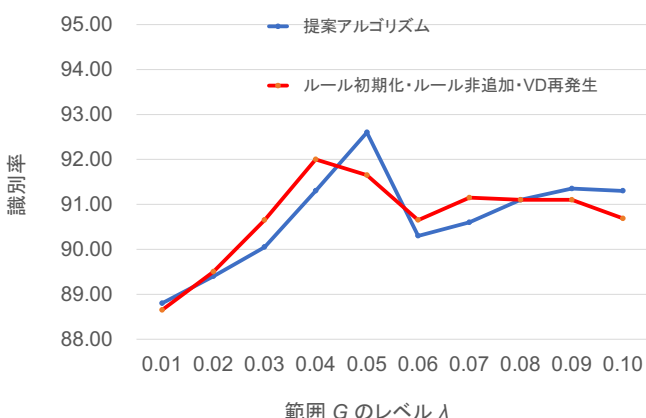


図 5: Discriminant Rate due to Size Change in Range G

5. おわりに

本論文では、多層間でバーチャルデータとファジィルールを継承する従来型の pdi-Boosting に加えて、ファジィルールの追加範囲 G を規則化した新たな pdi-BoostingG を定式化し、その有用性を議論した。特に、誤判別データが発生した場合の範囲 G の発生方法とその範囲 G 内のファジィルールの追加方法とバーチャルデータの発生方法について議論し、2 群識別の数値例により pdi-BoostingG の有用性を議論した。pdi-BoostingG は、範囲 G とバーチャルデータの発生により外乱に対するロバスト性に優れているが、今後、判別線の同定精度を向上させるため、より効率的な範囲 G の発生方法やバーチャルデータの発生法、サンプリング方法、継承率の調整法、ファジィルールの継承法について議論する必要がある。

なお、本研究の一部は、科学研究費補助金基盤研究 (C) 一般「バーチャルデータ発生型ファジィバギングを用いた放送映像の卓球戦略獲得ボードの開発」(No.20K11981, 2020 年～2025 年)、関西大学研究拠点形成支援経費「Harmonized Fitness: 音楽運動のアンサンブルによる健康づくりのスマート化」(2021 年～2022 年)、及び、2021 年

度関西大学学術研究員研究費の助成を得た。また、JST 次世代研究者挑戦的研究プログラム (JPMJSP2150) の支援を得た。

参考文献

- [1] 上田: アンサンブル学習, 情報処理学会論文誌, Vol.46, No.SIG15(CVIM12), pp.11-20 (2005)
- [2] R.Polikar: Ensemble Based Systems in Decision Making, *IEEE Circuits and Systems Magazine*, Vol.6, No.3, pp.21-45 (2006).
- [3] Y.Freund, R.E.Schapire: A Decision-theoretic Generalization of On-line Learning and An Application to Boosting, *Journal of Computer and System Sciences*, Vol.55, No.1, pp.119-139 (1997).
- [4] 三田: AdaBoost の基本原理と顔検出への応用, 情報処理学会研究報告, Vol.42(CVIM-159), pp.265-272 (2007)
- [5] 堀内, 大町, 阿曾: AdaBoost アルゴリズムを用いた識別手法の統合, 電子情報通信学会論文誌, Vol.J91-D, No.4, pp.1168-1171 (2008)
- [6] J.Friedman, T.Hastie, R.Tibshirani: Additive Logistic Regression: A Statistical View of Boosting, *Annals of Statistics*, Vol.28, No.2, pp.337-374 (2000).
- [7] A.Torralba, K.P.Murphy, W.T.Freeman: Sharing Features: Efficient Boosting Procedures for Multi-class Object Detection, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.762-769 (2004).
- [8] 高橋: クラス所属確率を用いた事例ごとの分類器選択, 言語処理学会第 15 回年次大会発表論文集, pp.709-712 (2009)
- [9] 安村, 上原: Bagging と Boosting を統合したアンサンブル学習方法, 第 19 回人工知能学会全国大会予稿集, No.3F1-01 (2005)
- [10] 林, 鶴背: 確率的データ補間を用いた BCI のための Boosting アルゴリズムの提案, 信学技報, Vol.109, No.461, pp.303-308 (2010)

- [11] I.Hayashi, S.Tsuruse, J.Suzuki, R.T.Kozma: A Proposal for Applying pdi-Boosting to Brain-Computer Interfaces, *Proceedings of 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE2012) in 2012 IEEE World Congress on Computational Intelligence (WCCI2012)*, pp.635-640 (2012).
- [12] 入江, 林: 正誤バーチャルデータの発生による pdi-Bagging の特性評価, 第 29 回インテリジェント・システム・シンポジウム講演論文集, Paper ID:No.A3-3 (2019)
- [13] 入江, 林: pdi-Bagging による発生バーチャルデータのクラス決定法の提案, 第 34 回人工知能学会全国大会講演論文集, Paper ID:No.103-GS-8-04 (2020)
- [14] 入江, 林, 堅多: pdi-Bagging を用いた大規模屋外駐車場で車種識別, インテリジェント・システム・シンポジウム 2021 講演論文集, pp.207-212 (2021)
- [15] 入江, 林: バーチャルデータとルールを継承するアンサンブル学習 pdi-Boosting の提案, 第 37 回ファジィシステムシンポジウム講演論文集, pp.118-123 (2021)
- [16] 市橋, 渡邊: 簡略ファジィ推論を用いたファジィモデルによる学習型制御, 日本ファジィ学会誌, Vol.2, No.3, pp.429-437 (1990)
- [17] 入江, 林: 台形型メンバシップ関数による学習型ファジィ推論の設計評価, 知能と情報, Vol.31, No.6, pp.908-917 (2019)

[連絡先]

林 勲 関西大学大学院 総合情報学研究科
〒 569-1095 大阪府高槻市霊仙寺町 2-1-1
tel. 072-690-2448
fax. 072-690-2491
e.mail ihaya@kansai-u.ac.jp