法政大学学術機関リポジトリ
HOSEI UNIVERSITY REPOSITORY

PDF issue: 2025-07-04

動径基底適応共鳴理論写像の分類機能について

岡本, 祐輔 / OKAMOTO, Yuusuke

(発行年 / Year)
2013-03-24
(学位授与年月日 / Date of Granted)
2013-03-24
(学位名 / Degree Name)
修士(工学)
(学位授与機関 / Degree Grantor)
法政大学 (Hosei University)

2012年度　修士論文

論文題名　動径基底適応共鳴理論写像の
分類機能について

指導教授　斎藤 利通 教授

法政大学大学院工学研究科
電気工学専攻修士課程
学生証番号： 11R3110
氏名　**岡本　祐輔**（オカモト　ユウスケ）

あらまし

適応共鳴理論ネットワーク (Adaptinve Resonance Theory nets, ART-nets) は、認知情報処理モデルのシステムである。自己組織化教師なし学習により、カテゴリの集合で分類を行う。この ART-nets を教師あり学習に拡張した適応共鳴理論写像 (Adaptinve Resonance Theory Map, ART-Map) が、提案されている。このシステムは、ビジランスパラメータを導入することで、入力を柔軟に学習することができる特徴がある。

本論文では、簡素な学習アルゴリズムとして提案されている動径基底 ARTMAP についての基本的な分類能力を考察する。特に入力順序が変化することによる分類能力の変化について 2 種類のベンチマーク問題を用いて検証考察する。

次に基本的な 2 層の RBART-net に中間層を導入した RBART-net とした、新しい分類アルゴリズムの階層型 RBART-Map を提案する。階層型に拡張することでカテゴリの統合を行うことを可能とした。また、RBART-net を 3 層にすることで、基本的な RBART-net では表現できない形を表すことができる。これを基本的なベンチマーク問題を用いて従来の RBART-Map と比較することで、この階層型 RBART-Map について検証する。

Analysis of the Classification Ability of
Radial Basis Adaptive Resonance Theory - Map

Abstract

The adaptive resonance theory networks (ART-Net) is unsupervised learning system and can form a categories to classify according to the input data on a feature space. The adaptive resonance theory Map (ART-Map) are supervised learning systems based on the ART-nets. This network have flexible classification functions.

In this paper, consider the basic classification ability of Radial Basis ART (RBART) using two benchmark problems. In particular, we analyze the classification capability when the input sequence changes.

Next, this paper a novel ART-Map with hierarchical structure and considers its classification function. Although existing ART-Maps have two-layer structure, the novel ART-Map has three-layer structure for more efficient and flexible classification function. The ART-Map can construct more various shapes of categories than that by existing ART-Maps. Performing numerical experiments for a typical benchmark, the algorithm efficiency is investigated.

目 次

第 1 章　まえがき **7**

1.1　Adaptive Resonance Theory 7

1.2　階層型 ART ネットワーク 9

第 2 章　動径基底 ART マップ **10**

2.1　学習アルゴリズム 10

2.2　数値実験 14

2.3　学習結果と基本特性の解析 15

第 3 章　階層型 RBART-Map **24**

3.1　階層型 RBART-Map 学習アルゴリズム 24

3.2　数値実験 29

3.3　エラー率 29

3.4　基本特性の解析 29

3.5　交差検証を用いた階層型 RBART-Map 35

3.5.1　交差検証 35

3.5.2　KFCV アルゴリズム 35

3.6　学習結果と比較 37

第 4 章　むすび **50**

参考文献 **51**

謝辞 **54**

図 目 次

1.1 代表的な ART-Map 7

1.2 ART ネットワークの構造 9

1.3 階層型 ART ネットワークの構造 9

2.1 初期化 12

2.2 維持 13

2.3 拡大 13

2.4 挿入 13

2.5 選択関数 14

2.6 エラーについて 15

2.7 入力 annulus 17

2.8 入力 fourclass 17

2.9 結果例 (入力 annulus) 18

2.10 結果例 (入力 fourclass) 18

2.11 エラー率 (入力 annulus) 19

2.12 エラー率 (入力 fourclass) 20

2.13 カテゴリ数 (入力 annulus) 21

2.14 カテゴリ数 (入力 fourclass) 22

2.15 入力順序が変化することによる分類能力が変化する考察１ 23

2.16 入力順序が変化することによる分類能力が変化する考察２ 23

3.1 維持 27

3.2 拡大 27

3.3 挿入 27
3.4 (a) 中間層に蓄えられたカテゴリ (b) 中間層のカテゴリの中心 ((a) 図中 ×) を入力として RBART を行なった結果 (c) 中間層のカテゴリを統合した結果 (出力層に蓄えられるカテゴリ) 28
3.5 樹形図 28
3.6 エラーについて 30
3.7 ρ_1, ρ_2 を変化させた時のエラー率 (入力 annulus) 31
3.8 ρ_1, ρ_2 を変化させた時のエラー率 (入力 forclass) 32
3.9 k_1, k_2 を変化させた時のエラー率 (入力 annulus) 33
3.10 k_1, k_2 を変化させた時のエラー率 (入力 fourclass) 34
3.11 入力を $k(k = 4)$ 個に分割する 36
3.12 k-分割交差検定 36
3.13 基本的な RBART-Map の結果例 (入力 annulus) 38
3.14 基本的な RBART-Map の結果例 (入力 fourclass) 38
3.15 階層型 RBART-Map の結果例 (入力 annulus) 39
3.16 階層型 RBART-Map の結果例 (入力 fourclass) 39
3.17 基本的な RBART-Map のエラー率 (入力 annulus) 40
3.18 階層型 RBART-Map のエラー率 (入力 annulus) 41
3.19 基本的な RBART-Map のカテゴリ数 (入力 annulus) 42
3.20 階層型 RBART-Map のカテゴリ数 (入力 annulus) 43
3.21 基本的な RBART-Map のエラー率 (入力 fourclass) 44
3.22 階層型 RBART-Map のエラー率 (入力 fourclass) 45
3.23 基本的な RBART-Map のカテゴリ数 (入力 fourclass) 46
3.24 階層型 RBART-Map のカテゴリ数 (入力 fourclass) 47
3.25 kFCV により決定した ρ_2 (入力 annulus) 48

3.26 基本的な RBART-Map と階層型 RBART-Map の分類能力の比較 (入力 annulus) 48
3.27 kFCV により決定した ρ_2 (入力 fourclass) 49
3.28 基本的な RBART-Map と階層型 RBART-Map の分類能力の比較 (入力 fourclass) 49

第1章　まえがき

1.1　Adaptive Resonance Theory

適応共鳴理論ネットワーク (Adaptinve Resonance Theory nets, ART-nets) は、S. Grossberg によって 1976 年に提案された認知情報処理モデルのシステムである。単位平方面の特徴空間上に入力された入力データを自己組織化教師なし学習により、カテゴリの集合で分類を行う [1]。この ART モデルの系列は、２値の入力パターンを学習する ART1[2] や連続な入力パターンに対応した Fuzzy ART-nets(FART-nets)[3] などが挙げられている。

また、この ART-nets を教師あり学習に拡張した適応共鳴理論写像 (Adaptinve Resonance Theory Map, ART-Map) が G. A. Carpenter, S. Grossberg, J. H. Reynolds 等によって 1991 年に提案されている [4]。この ART-map モデルの系列は、２値の入力パターンを学習するモデルから連続な入力パターンに対応したモデルが挙げられている。連続な入力パターンに対応した、代表的な ART-Map には図 1.1 に示すようにカテゴリが長方形で記述される Fuzzy ART-Map(FART-Map)[5]、楕円形で記述される Ellipisoid ART-Map(EART-Map)[6]、円形で記述される Hypersphere ART-Map(HART-Map)[7]、Radial BasisART-Map(RBART-Map)[8] [9] などが提案されている。

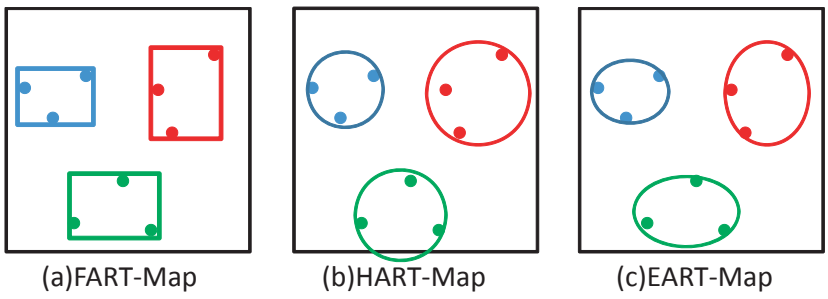


図 1.1: 代表的な ART-Map

これらのシステムは、入力データをカテゴリを用いて集合的に分類することで、簡素なアルゴリズムで学習することができ、高速な学習が可能であるという特徴がある。特に RBART-Map は、円形で記述される EART-Map に入力とカテゴリの類似度を測る新しい適合関数を導入することで、より簡素なアルゴリズムで学習を可能とした。また円形で記述されることで、カテゴリが多角形で記述される他の ART-Map より、メモリの消費量を抑えることができる特徴がある。これは、N 次元の入力データの時、超長方体で記述されるカテゴリのもつパラメータは $2N$ となるが、超球体で記述されるカテゴリは $N+1$ となり、次元が増えることでカテゴリがもつパラメータの数に膨大に差がでることから分かる。また実際に数値実験でも明らかにされている [8] [9]。

ART-Map の学習の特徴としてビジランスパラメータを導入することで、入力を柔軟に学習することを可能としている。ビジランスパラメータは閾値の役割をはたしており、入力とカテゴリの類似度 (適合関数) がこの値を超えると学習が行われる。基本的な ART-Map では、このパラメータを小さく設定するとカテゴリが粗く入力を学習するが、カテゴリ数は少なくなりメモリの消費量を減らすことができる。また、このパラメータの値を大きく設定すると、カテゴリが細かく正確に教師信号を学習するが、カテゴリ数が多くなりメモリの消費量は増える。この時 RBART-Map は、新しい適合関数を用いて分類を行なっているため、基本的な ART-Map とは反対に、ビジランスパラメータの値を小さく設定するとカテゴリが教師信号を細かく正確に分類し、この値を大きく設定するとカテゴリが粗く入力を学習する [4]。そのため、ビジランスパラメータを含むパラメータに対する ART-Map の学習機能の解析が盛んに行われている [8] [9]。現在では、パラメータを適切に設定することで、画像処理 [5] や信号処理 [10]、交通事故の解析 [11] と様々な応用も考えられてきている。また、巡回セールスマン問題を並列化して解くための並列化を行う学習アルゴリズムとしへの応用や [12]、論理合成への応用など様々な分野で発展する期待があり、これを確立する試みも研究されている。

第 2 章では、RBART-Map の基本的な学習アルゴリズムを示す。そしてこのシステムの入力データの入力順序を変化させることでこの分類機能に与える影響についていく

つかのベンチマーク問題を用いて解析し考察する。RBART-Map のパラメータにおける分類機能への影響は様々されてきたが、入力順序が変化することによりこの分類機能に与える影響は明らかにされていない。そのためこれを明らかにすることで RBART-Map の基本的な分類機能の解析のひとつになると考えこれを考察する。

1.2 階層型 ART ネットワーク

基本的な ART-Map は図 1.2 に示すように、2 層の ART ネットワークで構成されている [1]。入力層に蓄えられた入力データをもとにできたカテゴリが出力層に蓄えられる構造となっている。階層型 ART-Map は、図 1.3 に示すように 2 層のネットワークに中間層を挿入した 3 層のネットワークからなる ART-Map である。基本的な ART-Map では、パラメータの設定が適切であれば教師信号を正確に分類できるが、カテゴリ数が増大するおそれがある。中間層を挿入することで、入力層によりできるカテゴリを中間層に一度蓄えることでこのカテゴリを統合することを可能とした。また 3 層の ART-nets に拡張することで、2 層の ART-nets では表現できない形のカテゴリを表すことができる。これらのことにより基本的な ART-Map よりカテゴリ数を減らし教師信号を正確に分類することを可能とした。さらに基本的な ART-Map より汎化性の高い分類機能をもつ ART-Map となることを可能としている。

第 3 章では、この階層型 ART-Map を提案しアルゴリズムを示す。そしてこの階層型 ART-Map を RBART-Map に適応させ、2 つのベンチマーク問題を用いてこの分類機能の解析を行い考察する。

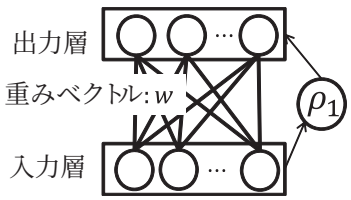


図 1.2: ART ネットワークの構造

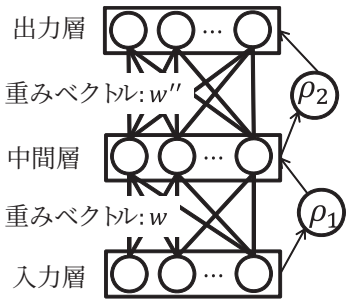


図 1.3: 階層型 ART ネットワークの構造

第2章　動径基底ARTマップ

2.1　学習アルゴリズム

図 1.2 に示す入力層には位置 P_i、クラス C_i によって特徴づけられる入力 I_i が N 個蓄えられる。

$$\boldsymbol{I}_i = (\boldsymbol{P}_i, C_i),\ \boldsymbol{P}_i \equiv (X_i, Y_i),\ C_i \in \{-1, 1\},\ i = 1 \sim N \tag{2.1}$$

出力層には離散時間 t における $M(t)$ 個のカテゴリを有し、各カテゴリは以下の重みベクトル W_j で特徴づけられる。この時、p_j はカテゴリの円の中心の位置、r_j はカテゴリの円の半径の大きさ、c_j はカテゴリのクラスを示す。

$$\boldsymbol{W}_j(t) = (p_j(t), r_j, c_j),\ p_j(t) \equiv (x_i(t), y_i(t)),\ \ c_j \in \{-1, 1\},\ j = 1 \sim M(t) \tag{2.2}$$

上記に記すように入力次元は 2 次元、クラスは 2 値に制限している。以下に学習アルゴリズムを示す。また、図に示す赤点はクラス +1 ($C_i = +1$) の入力、青点はクラス −1 ($C_i = -1$) の入力、赤丸はクラス +1 ($c_j = +1$) のカテゴリ、青丸はクラス −1 ($c_j = -1$) のカテゴリを表す。

Step 1: 初期化

離散時間 $t = 0$、$M(t) = 1$、ビジランスパラメータ $\rho_1 \in [0, \sqrt{2}]$、距離パラメータ $k_1 \in [-1, 1]$ とセットする。図 2.1(a) に示すように入力層に蓄えられた中からランダムに初期入力 $I_i(t)$ を提示する。そして以下の式に示した重みベクトル $w_1(t)$ を有するカテゴリを図 2.1(b) に示すように挿入する。この時、初期入力の位置 $P_i(t) \equiv (X_i(t), Y_i(t))$ を中心、半径を 0、カテゴリのクラスは初期入力のクラス C_i で表される。step2 へ。

$$\boldsymbol{W}_j(t) = (P_i(t), 0, C_i(t)),\ j = 1 \tag{2.3}$$

Step 2: 設定と入力

ρ_1 を初期化で設定した $\rho_1 \in [0, \sqrt{2}]$ に再度設定する。

離散時間 $t = t+1$ として入力層に蓄えられた入力 $I_i(t)$ をランダムに提示する (図 2.2(a) 参照)。step3 へ。

Step 3: 維持

図 2.2(a)(b) に示すように入力が既存のカテゴリ内に入力された時は、維持として、step10 へ。既存のカテゴリの円外に入力された時は、step4 へ。

Step 4: 勝者選択

$M(t)$ 個の各カテゴリについて、選択関数 T_j ($j = 1 \sim M(t)$) を計算する。$T_J = min(T_j)$ を有するカテゴリ W_J を勝者カテゴリとして、step5 へ。

$$\begin{aligned} T_j &= d_j - kr_j \\ d_j &= \sqrt{(x_j - X_i(t))^2 + (y_j - Y_i(t))^2} \end{aligned} \tag{2.4}$$

ここで d_j はカテゴリ W_j の中心と入力 $I_i(t)$ との距離である。また $k \in [-1, 1]$ は距離パラメータと呼ばれるパラメータであり、このパラメータの設定により選択関数 T_j は以下の図 2.5(a)(b)(c) のように変化する。

Step 5: クラスの比較

勝者カテゴリと入力のクラスの比較を行う。

・勝者カテゴリ W_j が有するクラスが入力 $I_i(t)$ のクラスと異なる場合、step6 へ。

・勝者カテゴリ W_j が有するクラスが入力 $I_i(t)$ のクラスと一致する場合、step7 へ。

・入力と一致するクラスのカテゴリが存在しないとき、step8 へ。

Step 6: マッチトラッキング

ビジランスパラメータ $\rho_1 = \rho_1 - \Delta\rho_1$ として、再度勝者選択を行う (step4 へ)。その際、離散時間 t において一度勝者カテゴリに選択された勝者カテゴリは”don’t care”とする。

Step 7: 行動選択

勝者カテゴリの有する選択関数 T_J を ρ_1 と比較を行う。

・$T_J \leq \rho_1$ の時、拡大として step8 へ。

・$T_J > \rho_1$ の時、維持として step9 へ。

Step 8: 拡大

勝者カテゴリと入力 $I_i(t)$ を含み図 2.3 のように拡大する。以下に勝者カテゴリの重みベクトルの更新式を示す。step10 へ。

$$
\begin{aligned}
&\boldsymbol{W}_J = (p_j^{(new)}, r_J^{(new)}, c_J),\\
&p_j^{new} \equiv (x_J^{(new)}, y_J^{(new)})\\
&x_J^{new} = \tfrac{1}{2}(r_J^{old}(\tfrac{x_J^{old} - X_i(t)}{d_J}) + x_J^{old} + X_i(t)),\\
&y_J^{new} = \tfrac{1}{2}(r_J^{old}(\tfrac{y_J^{old} - Y_i(t)}{d_J}) + y_J^{old} + Y_i(t)),\\
&r_J^{new} = \sqrt{(x_J^{new} - X_i(t))^2 + (y_J^{new} - Y_i(t))^2}.
\end{aligned}
\tag{2.5}
$$

Step 9: 挿入

$M(t) = M(t) + 1$ とする。そして図 2.4 のように入力 $I_i(t)$ を中心とした、半径 0、クラス $I_i(t)$ の新規カテゴリを挿入する。以下に新規カテゴリの重みベクトルの更新式を示す。

$$
\boldsymbol{W}_{M(t)} = (X_i(t), Y_i(t), 0, C_i(t)), \tag{2.6}
$$

Step 10: 終了判定

・$t < N$ の場合、step2 へ。

・$t = N$ の時終了。

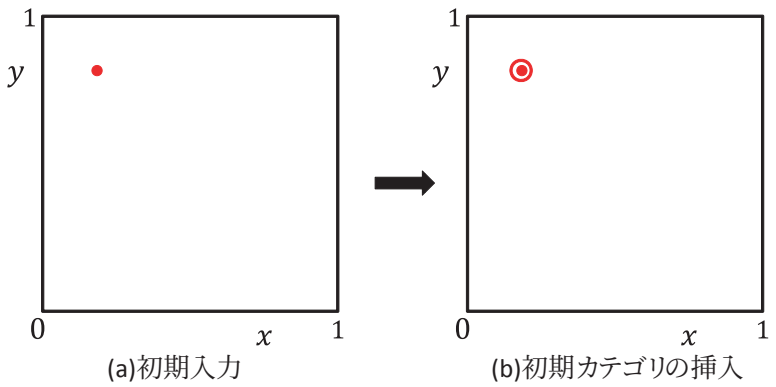


図 2.1: 初期化

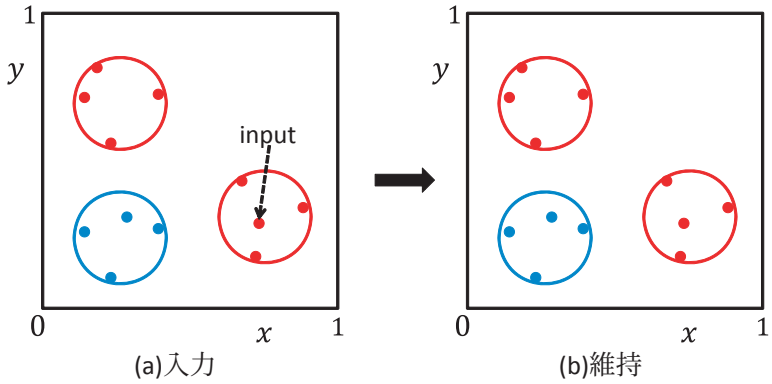


図 2.2: 維持

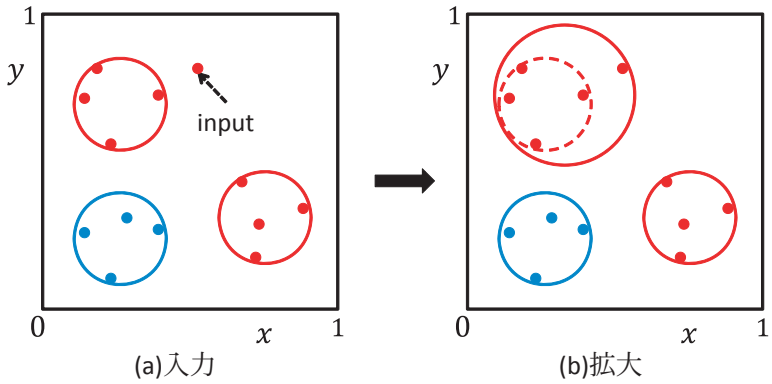


図 2.3: 拡大

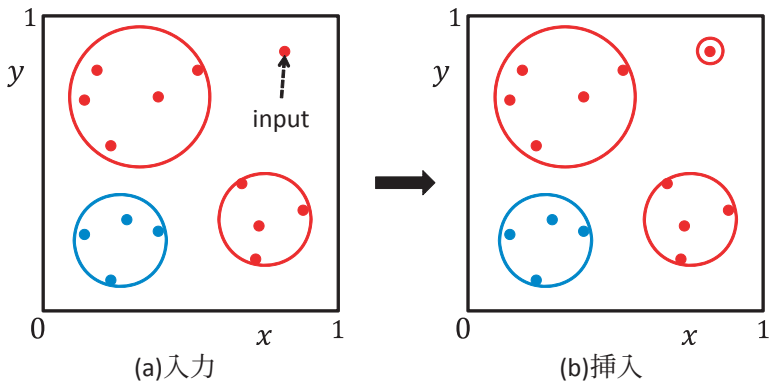
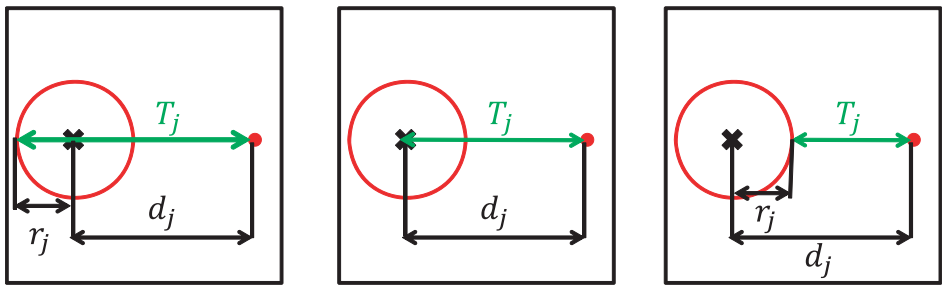


図 2.4: 挿入



(a) $k = -1$ (b) $k = 0$ (c) $k = +1$

図 2.5: 選択関数

2.2 数値実験

本章では、RBART-Map の基本的な分類能力について考察する。the University of California at Irvine（UCI）machine learning repository [15] が提供している annulus、fourclass の 2 種類のベンチマーク問題を用いて数値シミュレーションを行う。この 2 種類の入力の図を図 2.7, 図 2.8 に示した。青い丸 • はクラス –1 の入力、赤い四角 □ はクラス +1 の入力であり入力総数は 865 である。annulus、fourclass ともに 2 次元の 2 クラス問題である。annulus の入力のクラスは、中心 (0.5, 0.5) 半径 0.2 の円内と中心(0.5,0.5) 半径 0.4 の円外は –1 であり、それ以外のときは +1 である。fourclass は、4 クラス問題を 2 クラス問題に変換した問題である。非線形で分類しなければならないため 2 クラス問題の中でも複雑な問題とされている。この入力をすべて入力 (入力数 865) して学習を行い数値実験を行う。

これらの入力の入力順序をランダムに 50 回変化させ、RBART-Map の分類能力の変化を考察する。これにより入力順序が分類能力に与える影響について考察する。

この分類能力をはかるためにカテゴリ数とエラー率を算出する。カテゴリ数はカテゴリの数、エラー率は以下の式で定義する。

$$\text{エラー率} = \left(\frac{\text{エラー数}}{\text{入力総数}}\right)100 \tag{2.7}$$

エラー数とは入力のクラスとカテゴリのクラスが一致していない数であり、図 2.6 の × で示した入力がエラーである。この数がエラー数である。

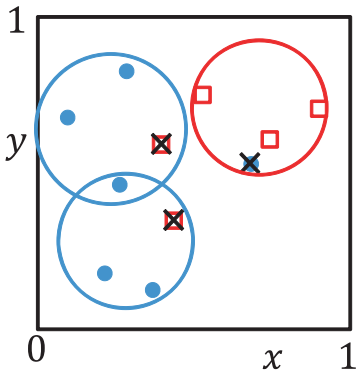


図 2.6: エラーについて

2.3　学習結果と基本特性の解析

学習結果例を図 2.9、図 2.10 に示す。● はクラス –1 の入力、赤い四角 □ はクラス +1 の入力である。この入力を分類している青い円はクラス –1 のカテゴリ、赤い円は +1 のカテゴリを表している。図 2.9 は annulus の分類問題の結果例である。図 2.10 は、fourclass の分類問題の結果例である。またビジランスパラメータ $\rho_1 = 0.2$、距離パラメータ $k_1 = 0$ に設定したときの結果例である。

入力順序を 50 回変化させた時の分類能力の結果を図 2.11~ 図 2.14 に示す。図 2.11、図 2.12 は、ビジランスパラメータ ρ_1 を 0.1 ~ 0.5 の範囲で 50 回入力順序をランダムに変化させたときの最大エラー率と最小エラー率を示す。この時の最大カテゴリ数と最小カテゴリ数を図 2.13、図 2.14 に示す。

図 2.11 の結果から ρ_1 が大きくなるにつれ最小エラー率、最大エラー率ともに上昇していることがわかる。また、最小エラー率、最大エラー率を比較するとエラー率に 20% 前後の差がうまれていることがわかる。

図 2.12 の結果から ρ_1 が大きくなると最大エラー率は上昇するが、最小エラー率は大きく上昇せずにほぼ横ばいの結果だとわかる。また、最小エラー率、最大エラー率を比較すると、ρ_1 が大きくなるにつれこの差が比例して大きくなっていることがわかる。

図 2.13、図 2.14 の結果から入力順序を変化させると、どちらの入力もカテゴリ数に差がうまれていることがわかる。また ρ_1 が大きくなるにつれ、この差が大きくなっていることがわかる。

これらの結果からRBART-Mapの基本的な分類能力について以下のことが述べられる。

- 入力順序を変化させるとエラー率、カテゴリ数が変化する。これは分類能力は入力順序に依存しているといえる。
- 入力順序を適切にすることでビジランスパラメータρ_1を大きくしても、エラー率を低くできる可能性がある。しかしこれは入力データに依存しており、すべての入力にいえるとは限らない。

ここでなぜ入力順序に依存するのか図2.15、図2.16の例を用いて説明する。図2.15、図2.16の入力の隣にある番号が入力順序をあらわしている。図2.15、図2.16では２番目と３番目の入力順序が入れ替わっているのみで、入力データは同じである。図2.15、図2.16を比較すると学習後のカテゴリが異なっている。これは図2.15においてカテゴリは２番目の入力で拡大しており３番目の入力でも拡大し、最終的に一つのカテゴリですべての分類していおりエラー数も多い。しかし、図2.16では２番目の入力で拡大せず挿入を行い入力を学習している。そして３番目の入力で拡大、４番目、５番目の入力で挿入の行動をとり入力を学習している。そのためカテゴリ数は多いがすべての入力を正確に分類している。このように入力順序が変化するとカテゴリの学習行動も変化することで分類能力に差がでていると考えられる。

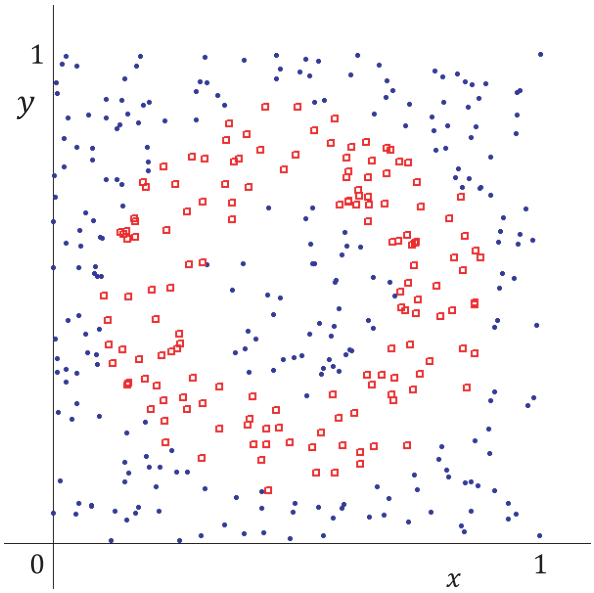


図 2.7: 入力　annulus

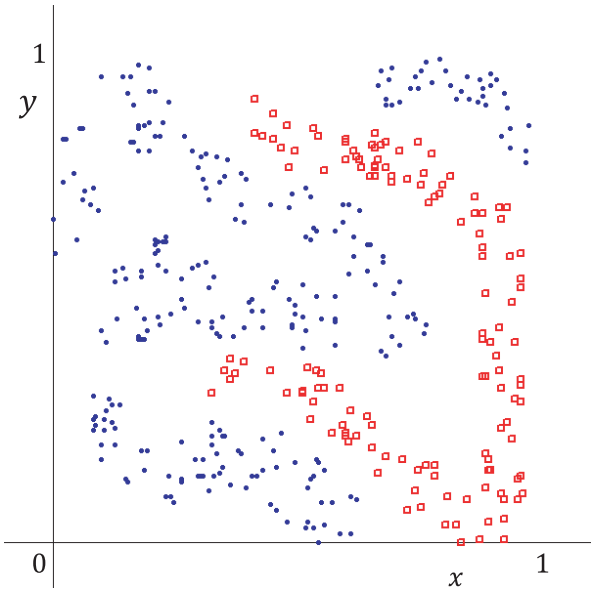


図 2.8: 入力　fourclass

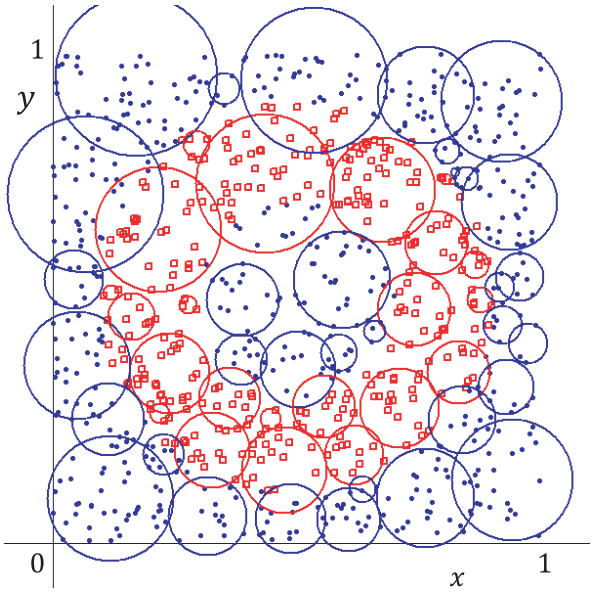


図 2.9: 結果例　(入力 annulus)

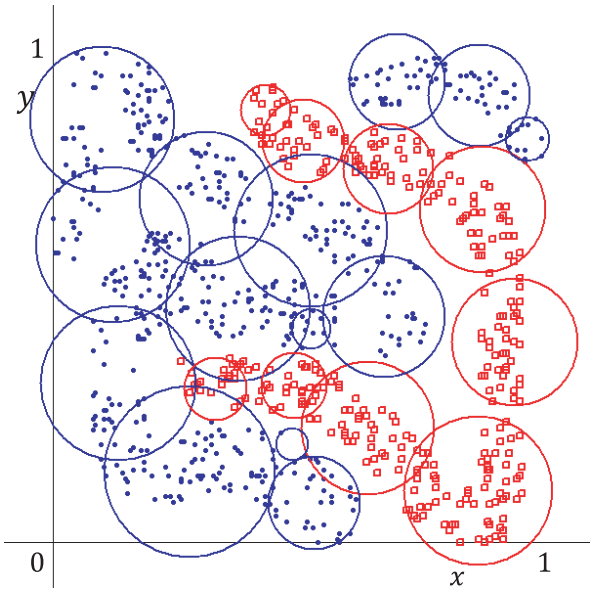


図 2.10: 結果例　(入力 fourclass)

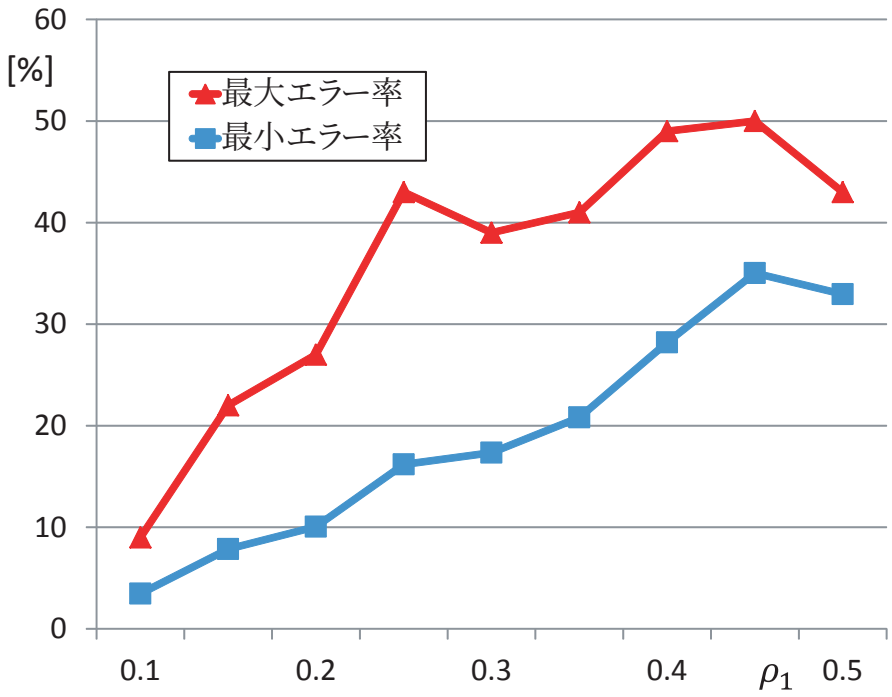


図 2.11: エラー率　(入力 annulus)

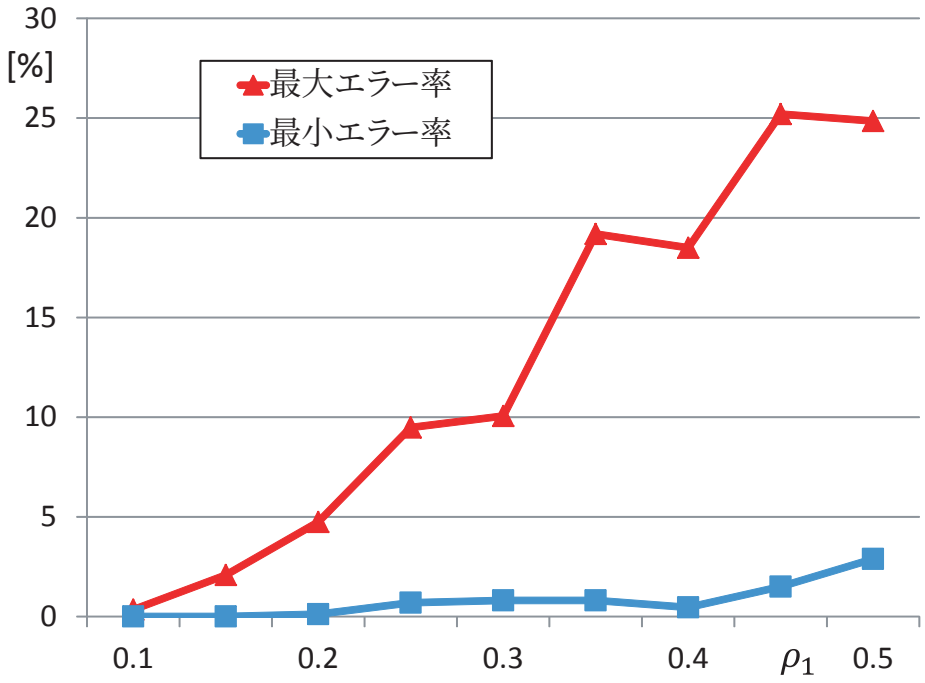


図 2.12: エラー率　(入力 fourclass)

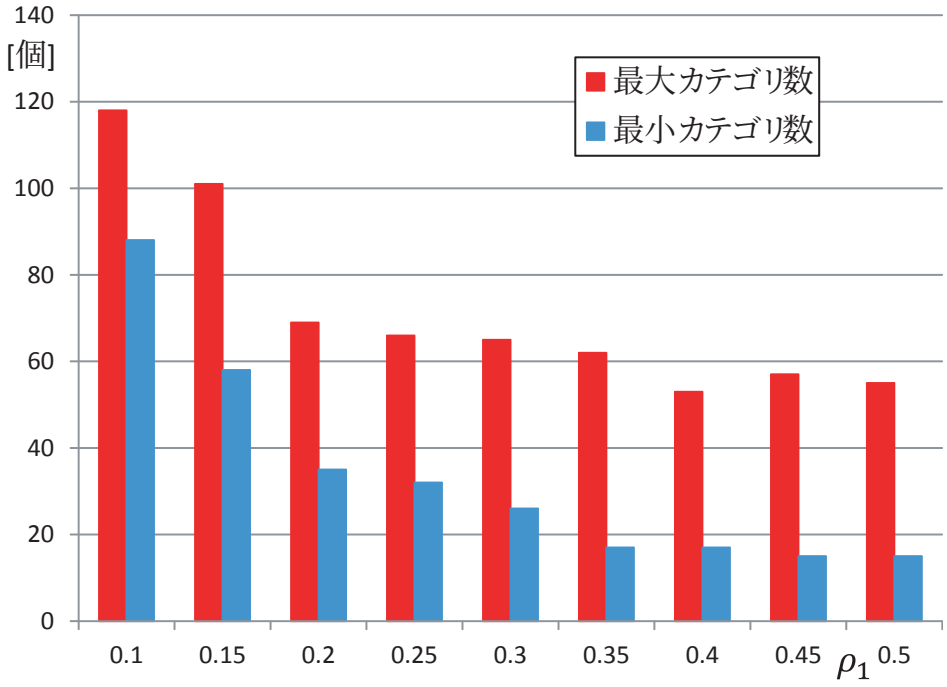


図 2.13: カテゴリ数　(入力 annulus)

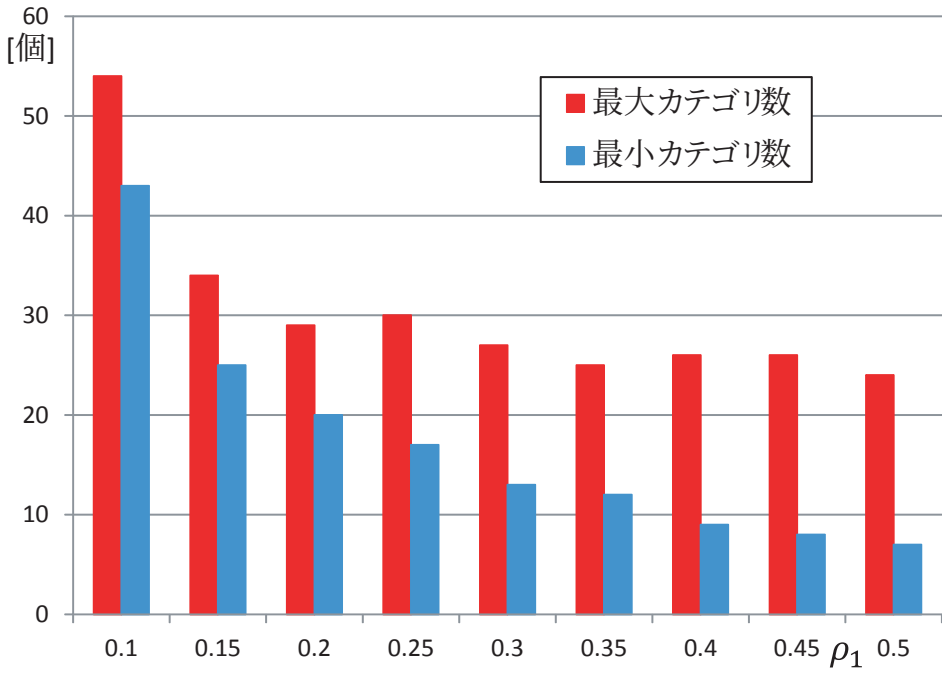


図 2.14: カテゴリ数　(入力 fourclass)

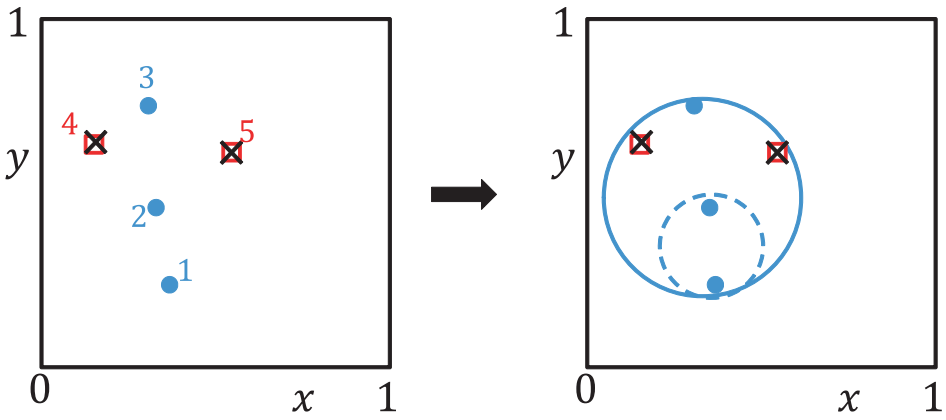


図 2.15: 入力順序が変化することによる分類能力が変化する考察 1

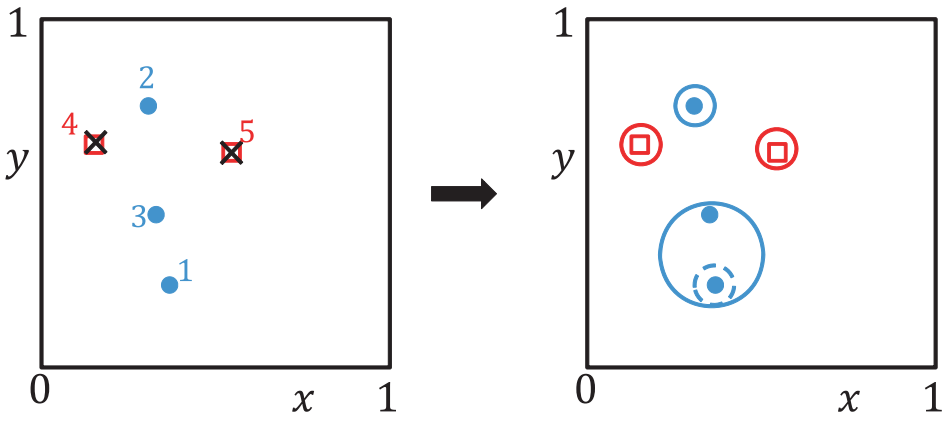


図 2.16: 入力順序が変化することによる分類能力が変化する考察 2

第3章　階層型RBART-Map

3.1　階層型RBART-Map学習アルゴリズム

図1.3に示すように入力層には位置P_i、クラスC_iによって特徴づけられる入力I_iがN個蓄えられる。

$$\boldsymbol{I}_i = (\boldsymbol{P}_i, C_i),\ \boldsymbol{P}_i \equiv (X_i, Y_i),\ C_i \in \{-1, 1\},\ i = 1 \sim N \tag{3.1}$$

中間層には離散時間tにおける$M(t)$個のカテゴリを有し、各カテゴリは以下の重みベクトルw_jで特徴づけられる。この時、p_jはカテゴリの円の中心の位置、r_jはカテゴリの半径の大きさを示す。また入力のクラスは2値に限定されているため、2クラス問題に特化した学習アルゴリズムで行う。これは、カテゴリが-1または$+1$のどちらかのクラスの入力を分類するアルゴリズムである。この時、そのクラスをc_bとする。

$$\boldsymbol{w}_j(t) = (p_j(t), r_j(t), c_b),\ p_j(t) \equiv (x_i(t), y_i(t)),\ \ c_b \in \{-1, 1\},\ j = 1 \sim M(t) \tag{3.2}$$

出力層には離散時間tにおける$S(t)$個のカテゴリを有し、中間層に蓄えられたカテゴリを統合するカテゴリw_s''が蓄えられる。また、図に示す赤点はクラス$+1$ ($C_i = +1$)の入力、青点はクラス-1 ($C_i = -1$)の入力、青丸はクラス-1 ($c_b = -1$)のカテゴリを表す。以下に、入力層から中間層にかけての学習アルゴリズムを示す。

Step 1: 初期化

離散時間$t = 0$、$M(t) = 1$、ビジランスパラメータ$\rho_1 \in [0, \sqrt{2}]$、距離パラメータ$k_1 \in [-1, 1]$、ベースクラス$c_b = \in [-1, +1]$とセットする。入力層に蓄えられた中からランダムに初期入力$I_i(t)$を提示する。そして以下の式に示した重みベクトル$w_1(t)$を有するカテゴリを挿入する。この時、初期入力の位置$P_i(t) \equiv (X_i(t), Y_i(t))$を中心、半径を0、

カテゴリのクラスは初期入力のクラス $C_i = c_b$ で表される。step2 へ。

$$\boldsymbol{w}_j(t) = (P_i(t), 0, c_b),\ j = 1 \tag{3.3}$$

Step 2: 入力

離散時間 $t = t+1$ として入力層に蓄えられた入力 $I_i(t)$ をランダムに提示する。step3 へ。

Step 3: ベースクラスの比較

入力が有するクラス $c_i(t) = c_b$ の時、step4 へ。さもなければ step8 へ。

Step 4: 維持

図 3.1(a) に示すように入力が既存のカテゴリ内に入力された時は、維持として、step8 へ。既存のカテゴリの円外に入力された時は、step5 へ。

Step 5: 勝者選択

$M(t)$ 個の各カテゴリについて、選択関数 T_j ($j = 1 \sim M(t)$) を計算する。$T_J = min(T_j)$ を有するカテゴリ w_J を勝者カテゴリとして、step6 へ。

$$\begin{aligned} T_j &= d_j - kr_j \\ d_j &= \sqrt{(x_j - X_i(t))^2 + (y_j - Y_i(t))^2} \end{aligned} \tag{3.4}$$

ここで d_j はカテゴリ w_j の中心と入力 $I_i(t)$ との距離である。また $k \in [-1, 1]$ は距離パラメータと呼ばれるパラメータであり、このパラメータの設定により選択関数 T_j は図 2.5(a)(b)(c) のように変化する。

Step 6: 行動選択

勝者カテゴリの有する選択関数 T_J を ρ_1 と比較を行う。

・$T_J \leq \rho_1$ の時、拡大として step7 へ。

・$T_J > \rho_1$ の時、維持として step8 へ。

Step 7: 拡大

勝者カテゴリと入力 $I_i(t)$ を含み図 3.2 のように拡大する。以下に勝者カテゴリの重み

ベクトルの更新式を示す。step8 へ。

$$
\begin{aligned}
&\boldsymbol{w}_J = (p_j^{(new)}, r_J^{(new)}, c_b),\\
&p_j^{new} \equiv (x_J^{(new)}, y_J^{(new)}\\
&x_J^{new} = \tfrac{1}{2}(r_J^{old}(\tfrac{x_J^{old}-X_i(t)}{d_J}) + x_J^{old} + X_i(t)),\\
&y_J^{new} = \tfrac{1}{2}(r_J^{old}(\tfrac{y_J^{old}-Y_i(t)}{d_J}) + y_J^{old} + Y_i(t)),\\
&r_J^{new} = \sqrt{(x_J^{new} - X_i(t))^2 + (y_J^{new} - Y_i(t))^2}.
\end{aligned}
\tag{3.5}
$$

Step 7: 挿入

M(t)=M(t)+1 とする。そして図 3.3 のように入力 $I_i(t)$ を中心とした、半径 0、クラス $I_i(t)$ の新規カテゴリを挿入する。以下に新規カテゴリの重みベクトルの更新式を示す。step8 へ。

$$
\boldsymbol{w}_{M(t)} = (X_i(t), Y_i(t), 0, c_b), \tag{3.6}
$$

Step 8: 終了判定

・$t < N$ の場合、step2 へ。

・$t = N$ の時 M(t)=M として、step9 へ。

以下からは中間層から出力層にかけてのカテゴリの統合を行う学習アルゴリズムである。

Step 9: 設定

離散時間 $t = 0$ とリセットし、ビジランスパラメータ $\rho_2 \in [0, \sqrt{2}]$、距離パラメータ $k_2 \in [-1, +1]$ を設定する。step10 へ。

Step 10: 統合するカテゴリの決定

中間層に蓄えられた、$\boldsymbol{w}_j,\ j = 1 \sim M$ の中心の位置を入力として、step2 から学習を行う。また、step8 の終了判定は $t = M$ となるまでこれを繰り返す判定とする。

Step 11: 統合

Step10 で中間層に蓄えられたカテゴリの中心が同一のカテゴリ $w_s', s = 1 \sim S$ で分類されたとき、そのカテゴリと、それに中心を含むカテゴリを統合する。図 3.4 の場合は次式のように重みベクトル w_s'' を更新しカテゴリを統合する。またこの統合の様子は図 3.5 のように樹形図で表すことができる。

$$
\begin{aligned}
w_1'' &= (w_1 \cup w_2) \cup w_1'\\
w_2'' &= w_3 \cup w_2'
\end{aligned}
\tag{3.7}
$$

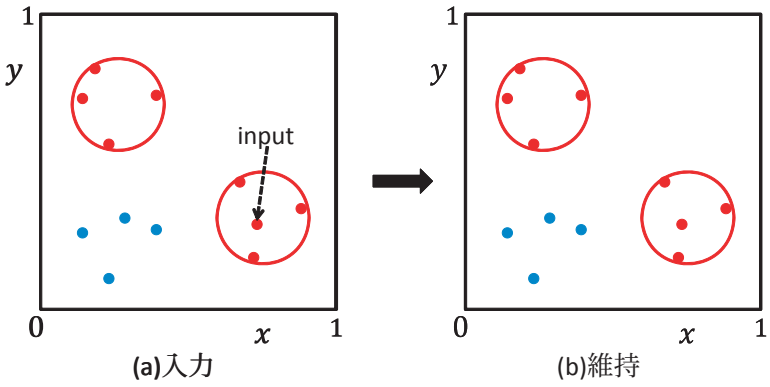


図 3.1: 維持

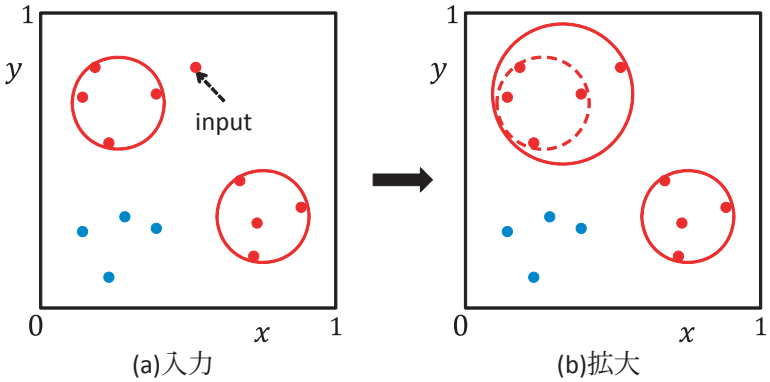


図 3.2: 拡大

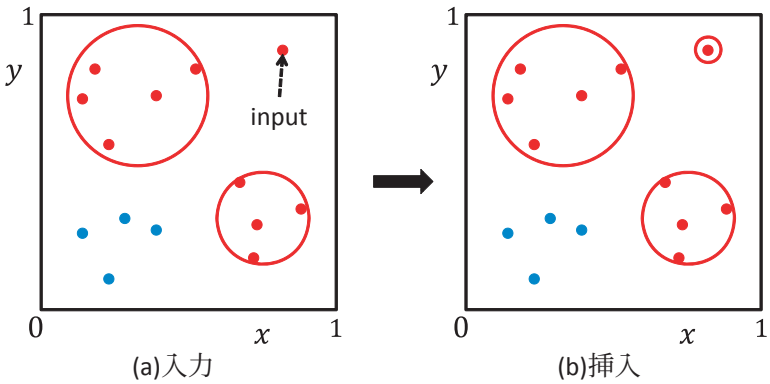


図 3.3: 挿入

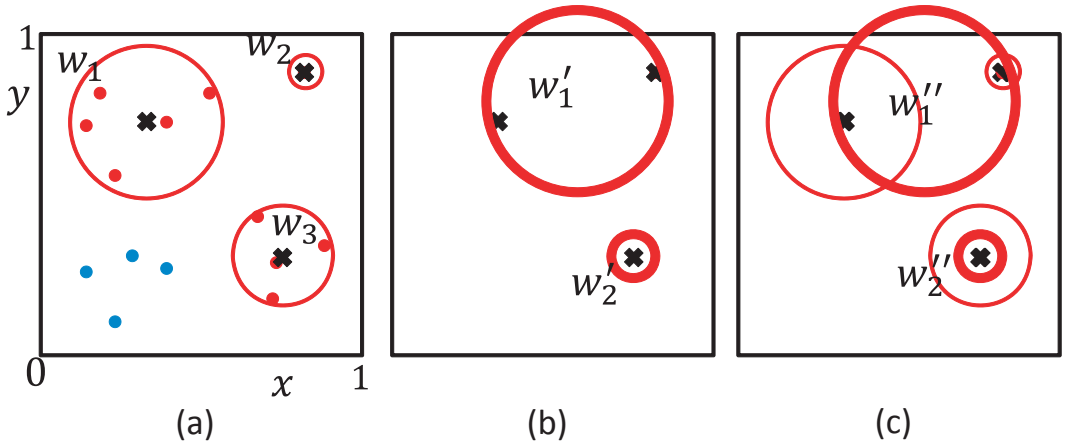


図 3.4: (a) 中間層に蓄えられたカテゴリ (b) 中間層のカテゴリの中心 ((a) 図中 ×) を入力として RBART を行なった結果 (c) 中間層のカテゴリを統合した結果 (出力層に蓄えられるカテゴリ)

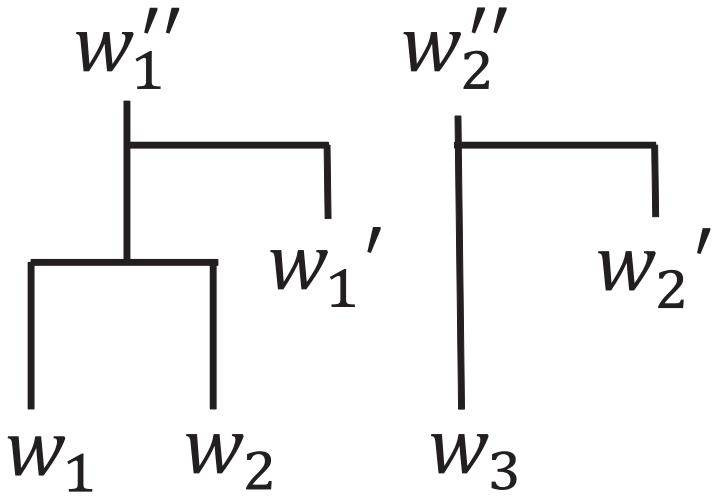


図 3.5: 樹形図

3.2　数値実験

階層型 RBART-Map の基本的な分類能力を考察する。そのために the University of California at Irvine（UCI）machine learning repository [15] が提供している２つの基本的な分類問題を用いて行う。これは図 2.9、図 2.10 に示すように、２次元の入力データである。またクラスは 2 値であり、赤い四角 ロはクラス+1 の入力。青い丸 ● はクラス-1 の入力である。ともに入力総数は 865 である。この時学習のために用いる訓練データはランダムに $N = 400$ 個選択する。

3.3　エラー率

ここで、分類能力を考察するためにエラー率を定義する。エラー率とは、学習後のカテゴリが教師信号を正確に分類できていない割合である。この時、カテゴリはどちらかのクラス c_b のみを分類するため、図 3.6 の × で示すようにカテゴリに属していない c_b クラスの入力と、異なるカテゴリで分類された入力をエラーとする。そして以下の式でエラー率を定義する。

$$\text{エラー率} = \left(\frac{\text{エラー数}}{\text{入力総数}}\right)100 \tag{3.8}$$

3.4　基本特性の解析

階層型 RBART-Map の基本的な分類能力を考察するために、ビジタランスパラメータ ρ_1, ρ_2 を 0.1 ~ 0.5 の範囲でそれぞれ変化させた時のエラー率の結果を図 3.7、図 3.8 に示す。図 3.7 は入力 annulus の時のエラー率を示しており、図 3.8 は入力が fourclass の時のエラー率の結果である。また距離パラメータ k_1, k_2 ともに 0 である。次に k_1, k_2 を -1 ~ $+1$ の範囲でそれぞれ変化させた時の結果を図 3.9、図 3.10 に示す。図 3.9 は入力が annulus の時のエラー率、図 3.10 は入力が fourclass の時のエラー率である。この時ビジランスパラメータ ρ_1, ρ_2 ともに 0.2 である。

図 3.7、図 3.8 の結果からビジランスパラメータ ρ_1 が小さいとき、ρ_2 が変化するとエラー率が大きく変化しており、ρ_2 を適切に設定することでエラー率を低くすること

ができる可能性があることがわかる。ρ_1 が大きい時、ρ_2 の設定に関わらず、あまりエラー率に影響しないことがわかる。これらの理由として、ρ_1 を大きく設定すると入力層によりできる中間層に蓄えられるカテゴリの大きさが大きく、このカテゴリが統合せずそのまま出力層に蓄えられていることが考えられる。逆に ρ_1 を小さく設定すると中間層に蓄えられるカテゴリの大きさが小さくなり、このカテゴリを統合しやすくなるため、ρ_2 の設定によりエラー率が大きく変化すると考えられる。

距離パラメータも同様に図 3.9、図 3.10 の結果から k_1 が小さいとき k_2 が変化するとエラー率が大きく変化していることがわかる。k_1 が大きい時、k_2 の設定に関わらず、あまりエラー率に影響しないことがわかる。これらの理由もビジランスパラメータと同等な理由が考えられる。

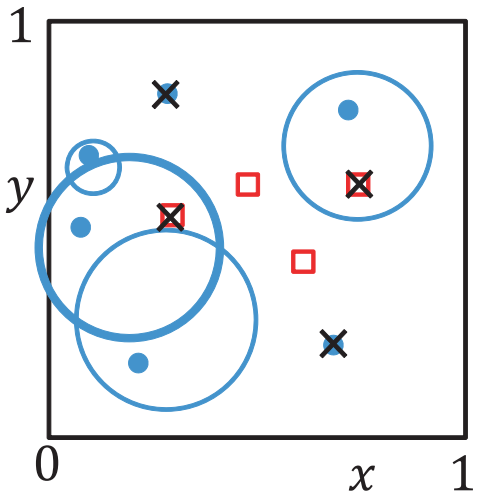


図 3.6: エラーについて

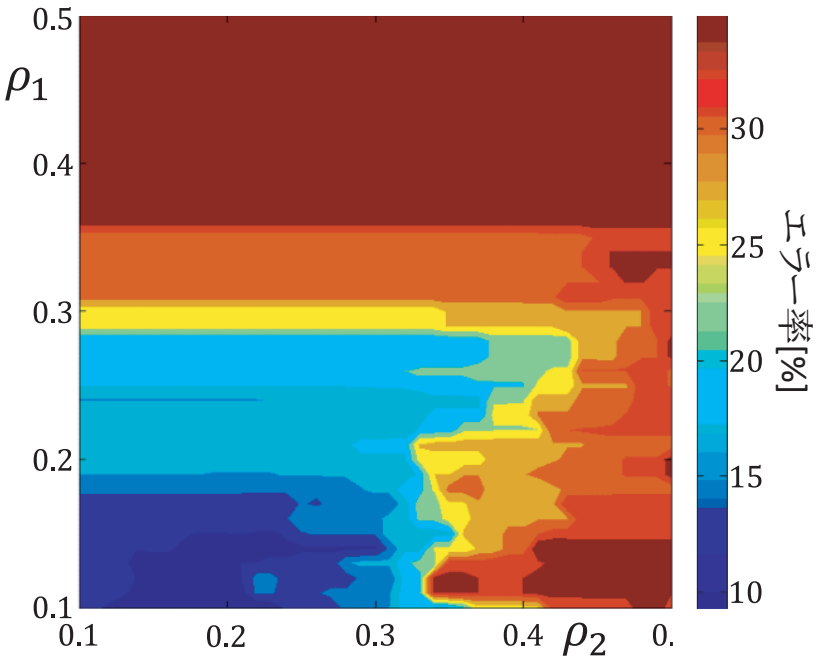


図 3.7: ρ_1, ρ_2 を変化させた時のエラー率 (入力 annulus)

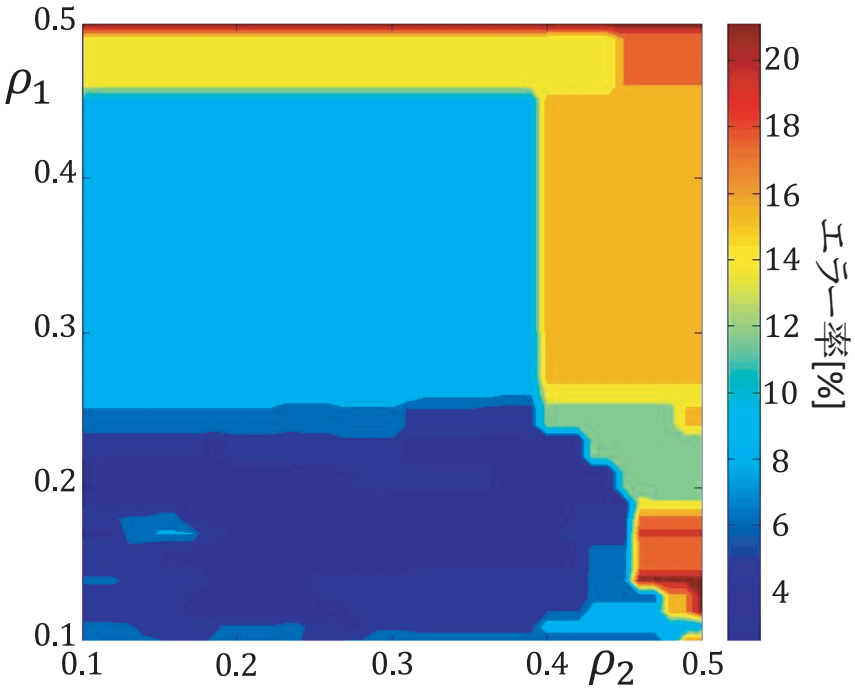


図 3.8: ρ_1, ρ_2 を変化させた時のエラー率 (入力 forclass)

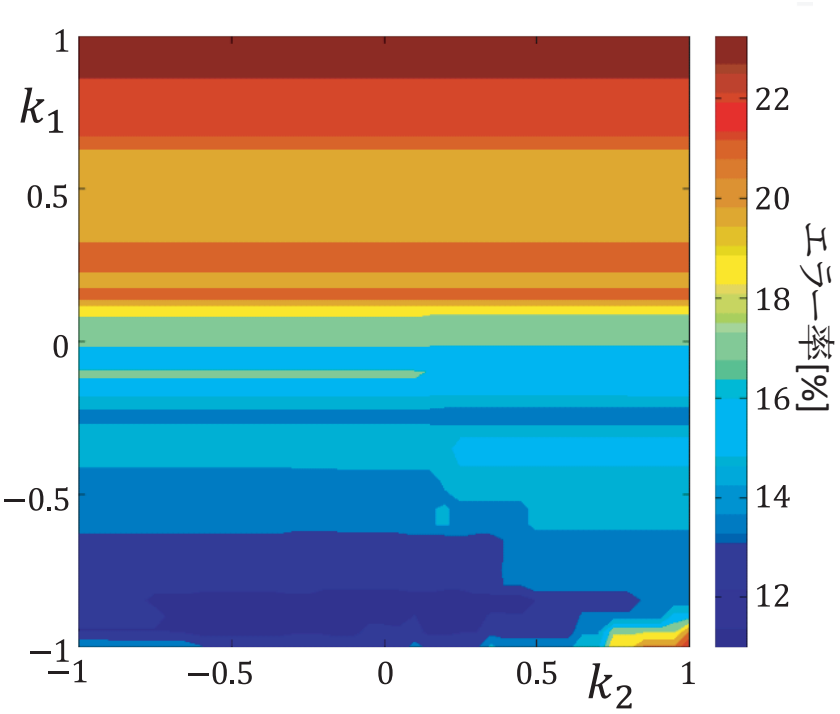


図 3.9: k_1, k_2 を変化させた時のエラー率 (入力 annulus)

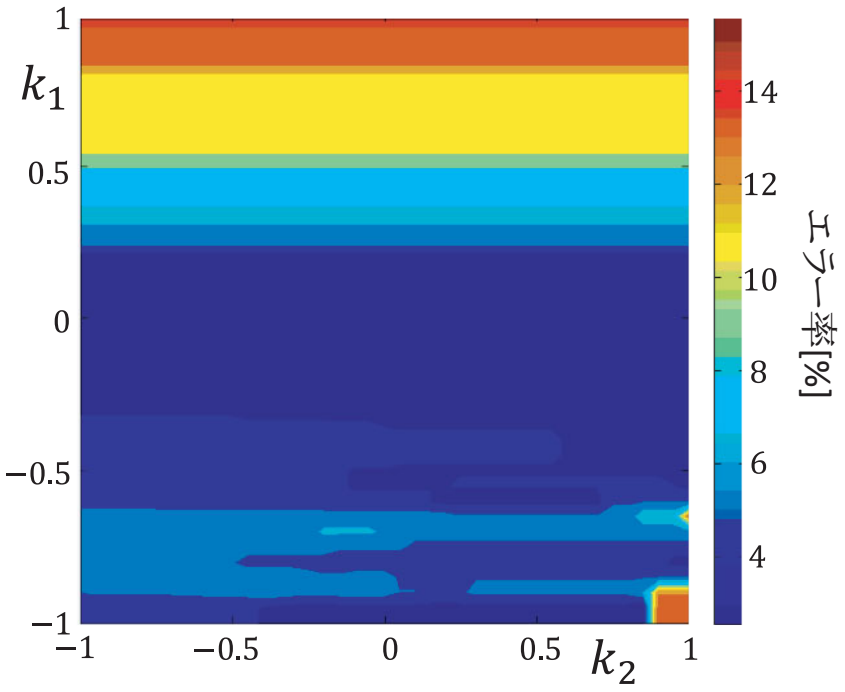


図 3.10: k_1, k_2 を変化させた時のエラー率 (入力 fourclass)

3.5　交差検証を用いた階層型 RBART-Map

3.5.1　交差検証

交差検証 (Cross Validation, CV) は、モデルのパラメータを決定することや様々なモデルからどのモデルが最適なのかを考えることができる手法である [13]。分類モデルにおいて汎化能力の高いモデルにすることは重要である。既知のデータが十分にある際には、この既知のデータの一部を用いて汎化能力の高いモデルを選択することができる。しかし、基本的に既知のデータは限られており、良いモデルを作るには既知のデータを訓練にできるだけ多く使用することが望ましい。しかし検証用データが少ないと汎化性が低くなるというジレンマが生じる。このジレンマを解消するために CV が提案された。代表的なモデルには、leave-one-out CV、k-分割交差検定な (K-fold cross-validation, KFCV) どが挙げられている [13][14]。

本論文では KFCV を階層方 RBART-Map のパラメータ ρ_2 の設定に応用する。階層型 RBART-Map は基本的特性でも述べたように、ρ_1 が小さいとき ρ_2 を適切に設定することで汎化性の高い分類を行うことができる。そのため、KFCV を利用することで ρ_1 が決まれば自動的に汎化能力の高い ρ_2 のパラメータに設定することを考える。

この結果を 2 値分類に特化した基本的な RBART-Map の結果と比較して有効性を確認する。

3.5.2　KFCV アルゴリズム

入力 I_i を図 3.11 に示すように $k(k = 4)$ 個に分割する。

図 3.12 に示すように、そのうちの 1 つを検証データとし，残る $k – 1$ 個を訓練データとし、これを $k(k = 4)$ 回行う。この時の評価値の平均値をとることで推定を行う。

本論文では、訓練データを学習に用いるデータとし、検証用データをこの学習後のカテゴリに入力することで評価値 (エラー率) を算出する。そしてビジランスパラメータ ρ_1 を 0.1 ~ 0.5 の範囲で変化させ上記の KFCV を行い、一番平均エラー率の低い ρ_2 に決定する。

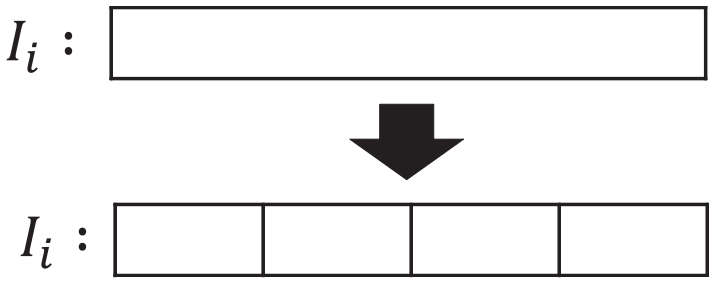


図 3.11: 入力を $k(k = 4)$ 個に分割する

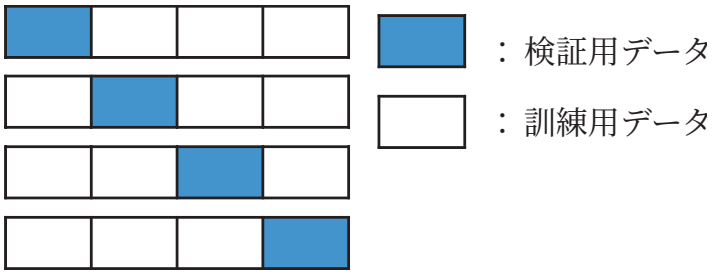


図 3.12: k-分割交差検定

3.6　学習結果と比較

基本的な RBART-Map の結果例を図 3.13、図 3.14 に示す。この時パラメータはそれぞれ$\rho_1 = 0.2$、$k_1 = 0$である。交差検証を用いた階層型 RBART-Map の結果例を図 3.15、図 3.16 に示す。図 3.15 のパラメータはそれぞれ$\rho_1 = 0.2$、$\rho_2 = 0.27$、$k_1 = 0, k_2 = 0$である。図 3.16 のパラメータはそれぞれ$\rho_1 = 0.2$、$\rho_2 = 0.37$、$k_1 = 0, k_2 = 0$である。またこの時、3.15、図 3.16 の太い線の青丸は細い線の青丸を統合していることを示している。学習に用いる入力数$N = 400$とする。これを学習後にこの入力を含むすべての入力 (入力数 865) を入力して分類能力を測る。距離パラメータk_1、k_2はすべて 0 に設定した結果である。

annulus の入力を行なったときの基本的な RBART-Map の分類能力の結果と交差検証を用いた階層型 RBART-Map の分類能力の結果を、図 3.18 ~ 図 3.20 に示す。fourclass の入力を行なったときの基本的な RBART-Map の分類能力の結果と交差検証を用いた階層型 RBART-Map の分類能力の結果を、図 3.21 ~ 図 3.24 に示す。これらの結果を比較した結果をに示す。図 3.26 は annulus の入力を行なったときの比較結果である。図 3.28 は fourclass の入力を行なったときの比較結果である。この時 KFCV により決定したrho_2の値を図 3.25、図 3.27 に示す。

これらの結果から以下のことがわかる。

- 階層型 RBART-Map は、ρ_1が小さいときρ_2を KFCV を用いて適切に設定することでエラー率も少なく、基本的な RBART-Map よりもカテゴリ数を大幅に減少させることを可能とした。またこの結果からρ_1が小さいとき KFCV を用いてρ_2をほとんど適切に設定できていることが分かる。

- ρ_1が大きいとき、エラー率、カテゴリ数ともに階層型 RBART-Map と基本的な RBART-Map ではあまり変わらないことがわかる。これは階層型 RBART-Map の基本的特性でも述べたように、ρ_1が大きい時中間層に蓄えられたカテゴリは統合されないため、基本的な RBART-Map とは変わらない結果となる。

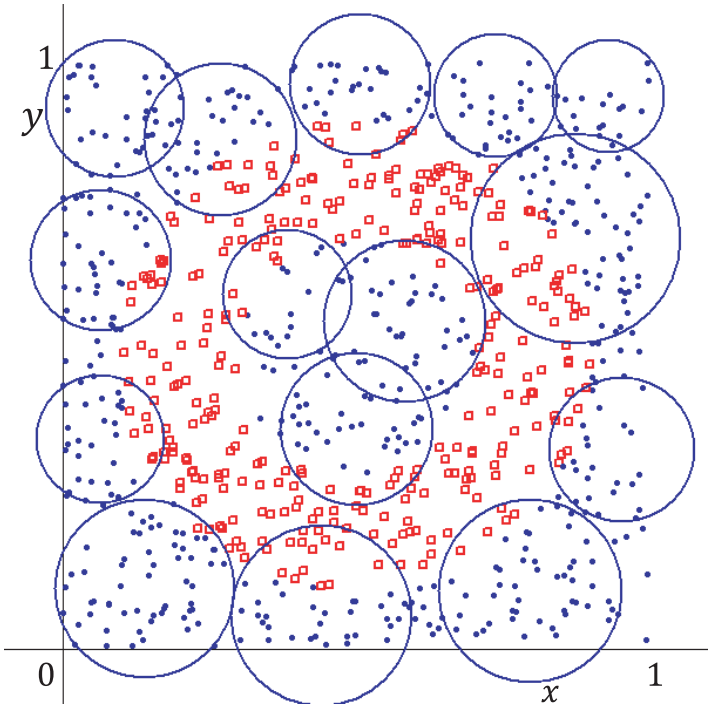


図 3.13: 基本的な RBART-Map の結果例 (入力 annulus)

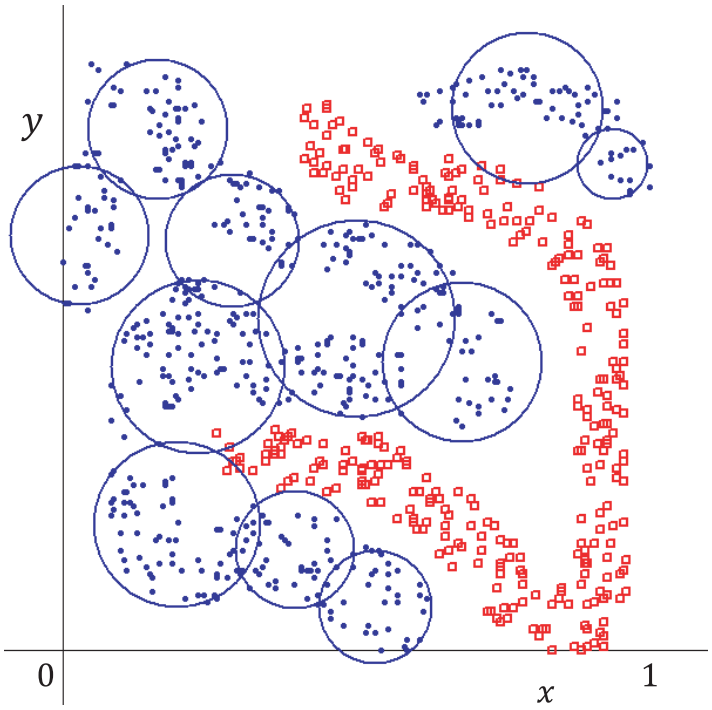


図 3.14: 基本的な RBART-Map の結果例 (入力 fourclass)

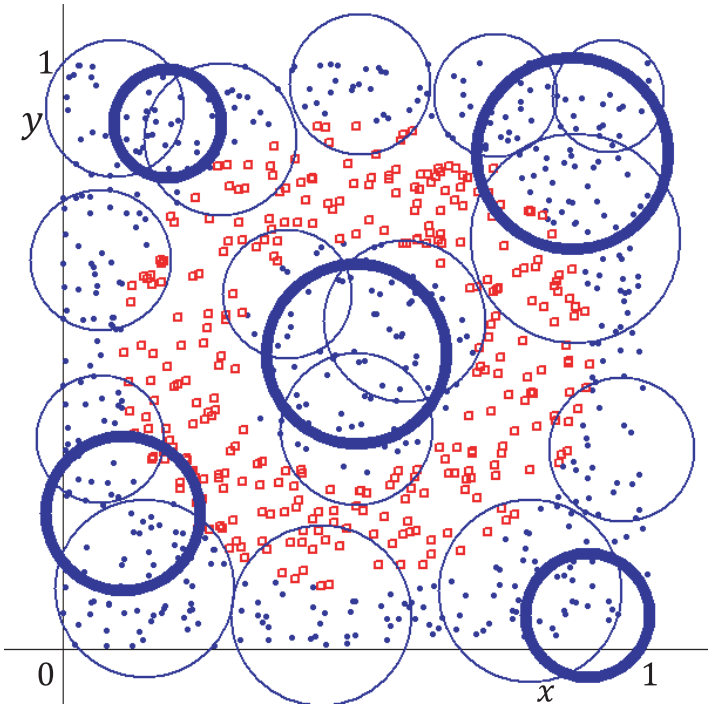


図 3.15: 階層型 RBART-Map の結果例 (入力 annulus)

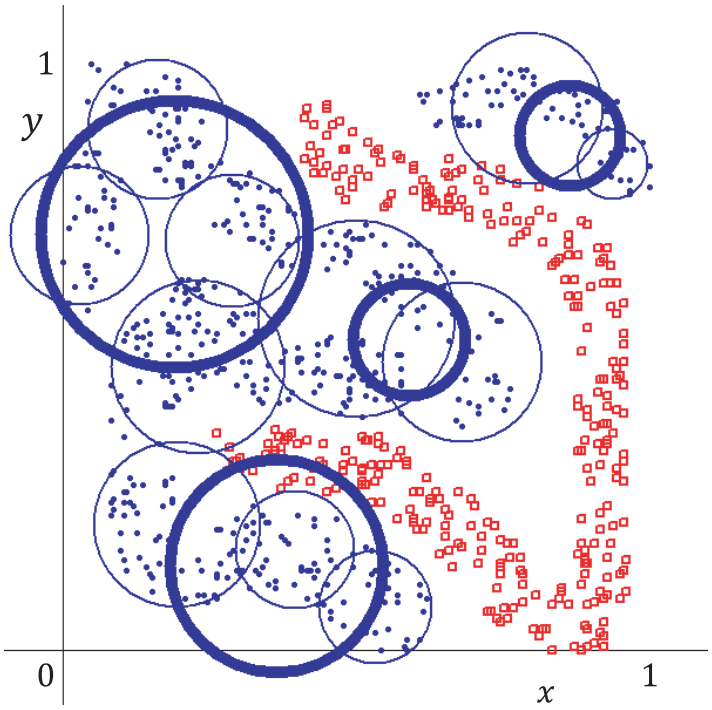


図 3.16: 階層型 RBART-Map の結果例 (入力 fourclass)

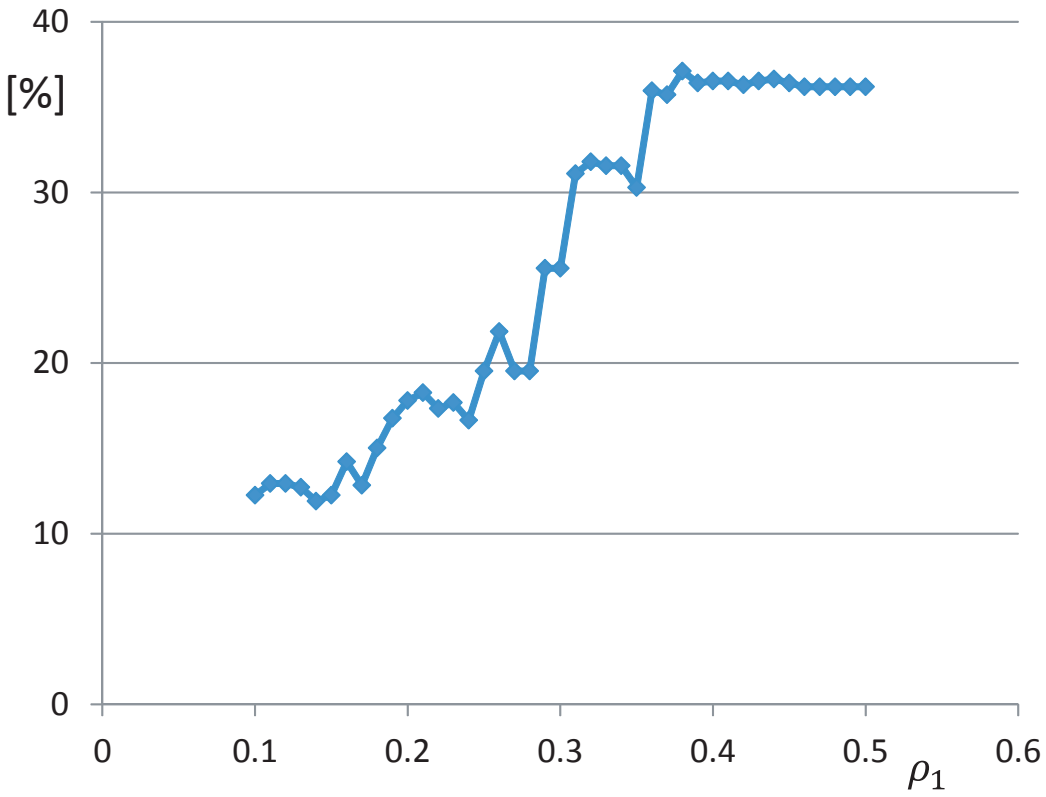


図 3.17: 基本的な RBART-Map のエラー率 (入力 annulus)

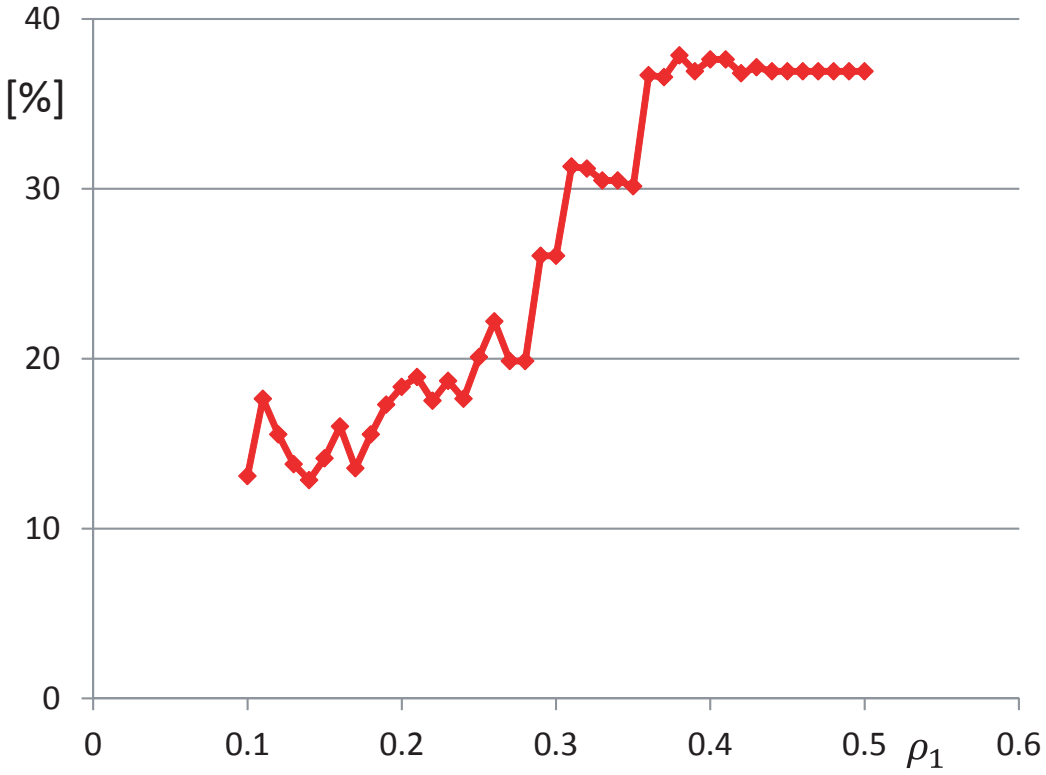


図 3.18: 階層型 RBART-Map のエラー率 (入力 annulus)

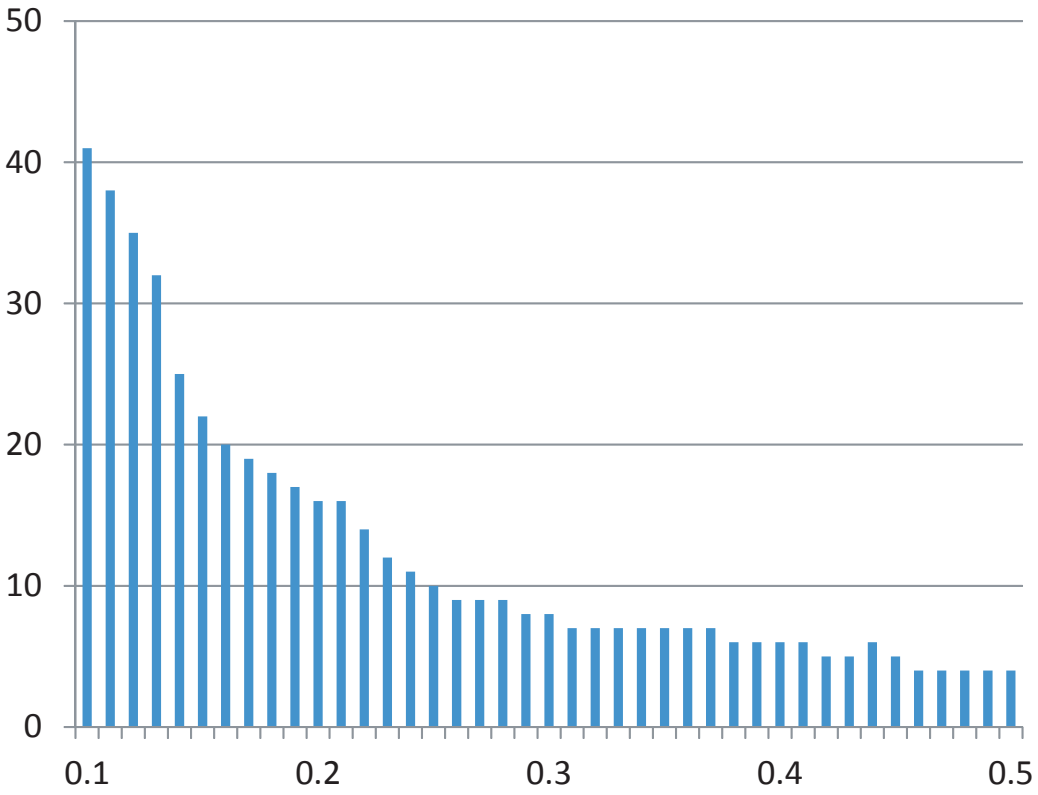


図 3.19: 基本的な RBART-Map のカテゴリ数 (入力 annulus)

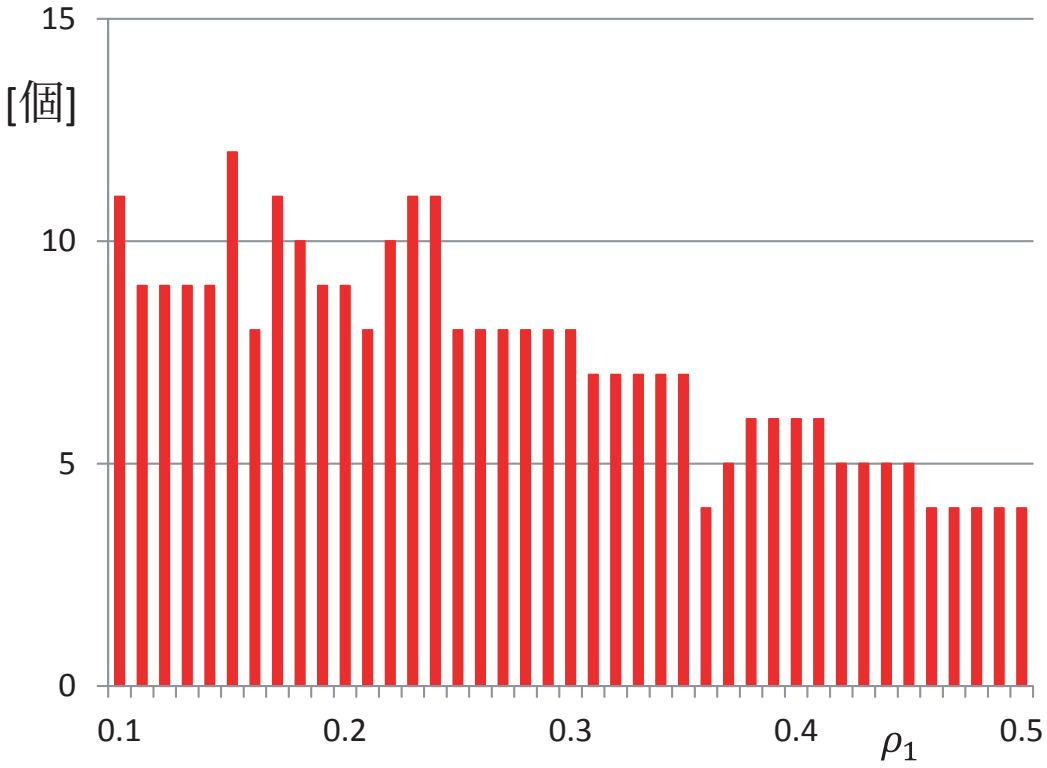


図 3.20: 階層型 RBART-Map のカテゴリ数 (入力 annulus)

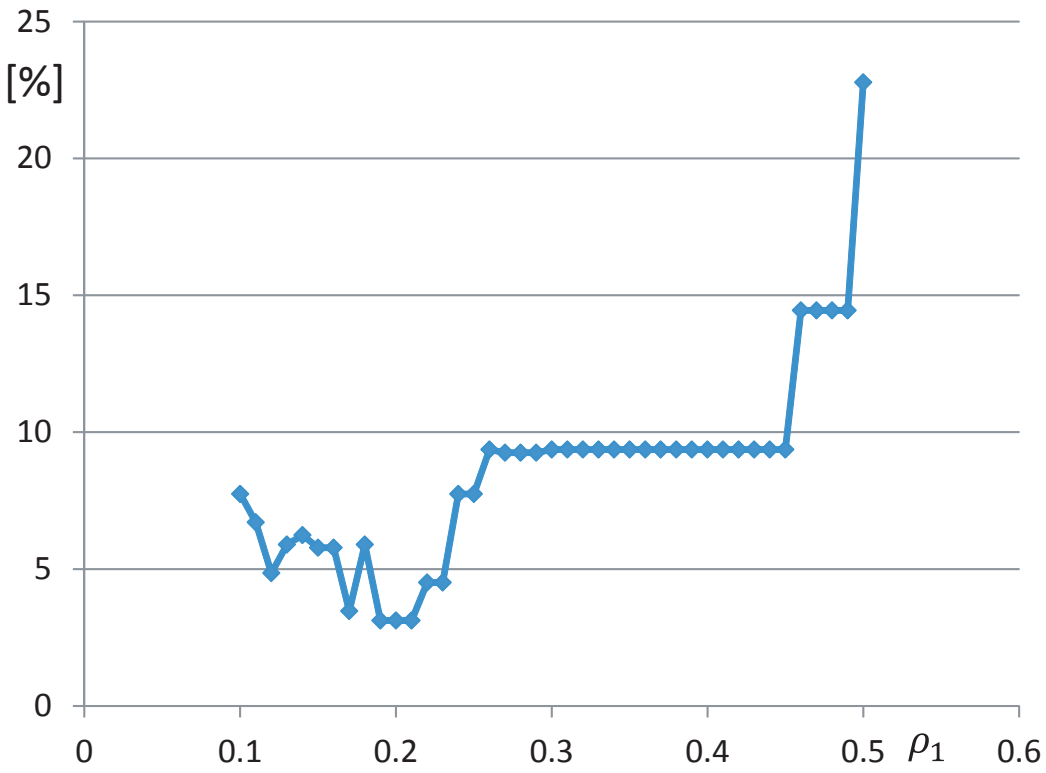


図 3.21: 基本的な RBART-Map のエラー率 (入力 fourclass)

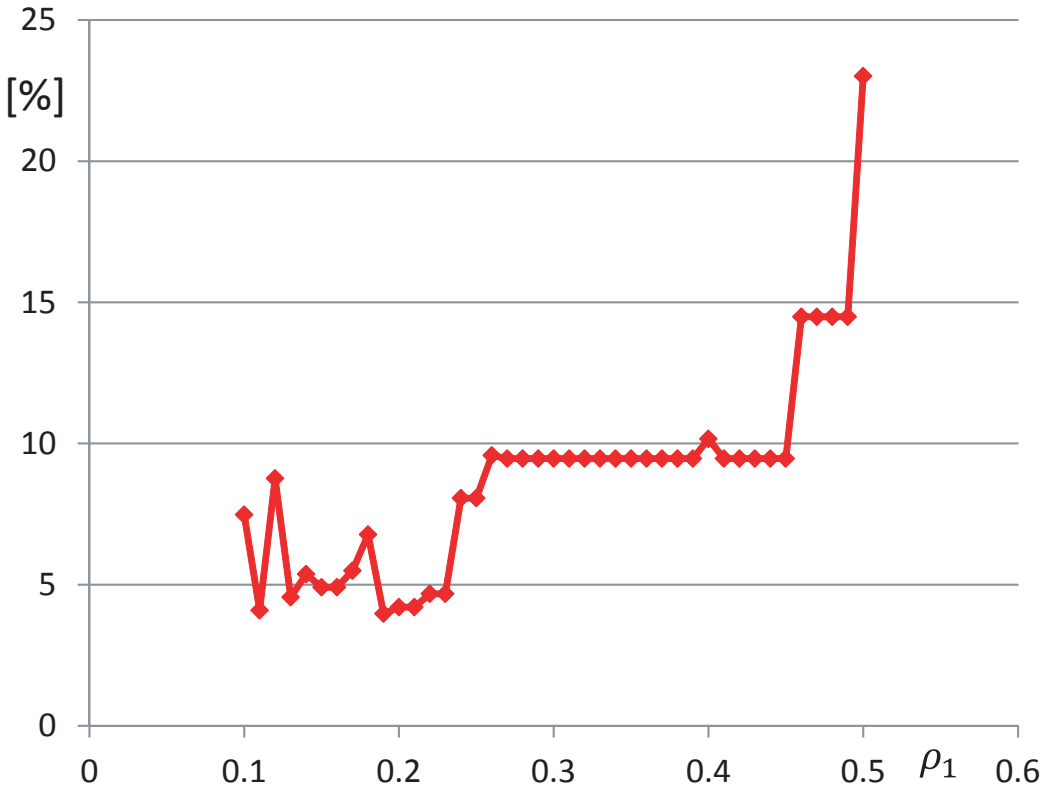


図 3.22: 階層型 RBART-Map のエラー率 (入力 fourclass)

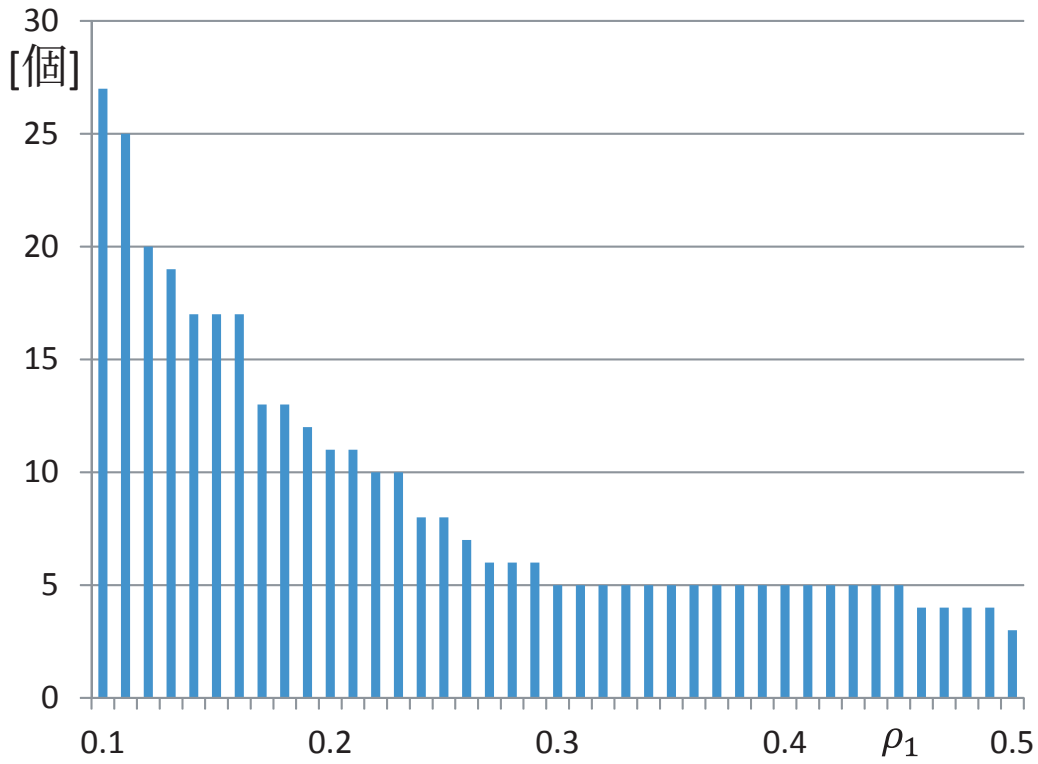


図 3.23: 基本的な RBART-Map のカテゴリ数 (入力 fourclass)

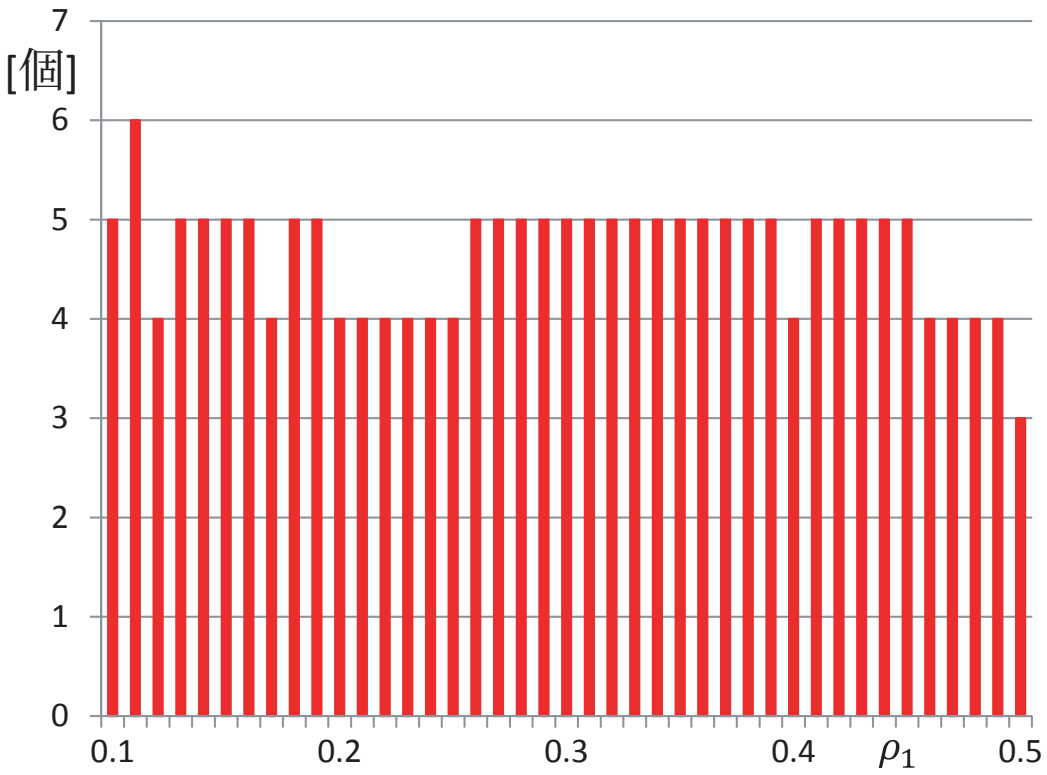


図 3.24: 階層型 RBART-Map のカテゴリ数 (入力 fourclass)

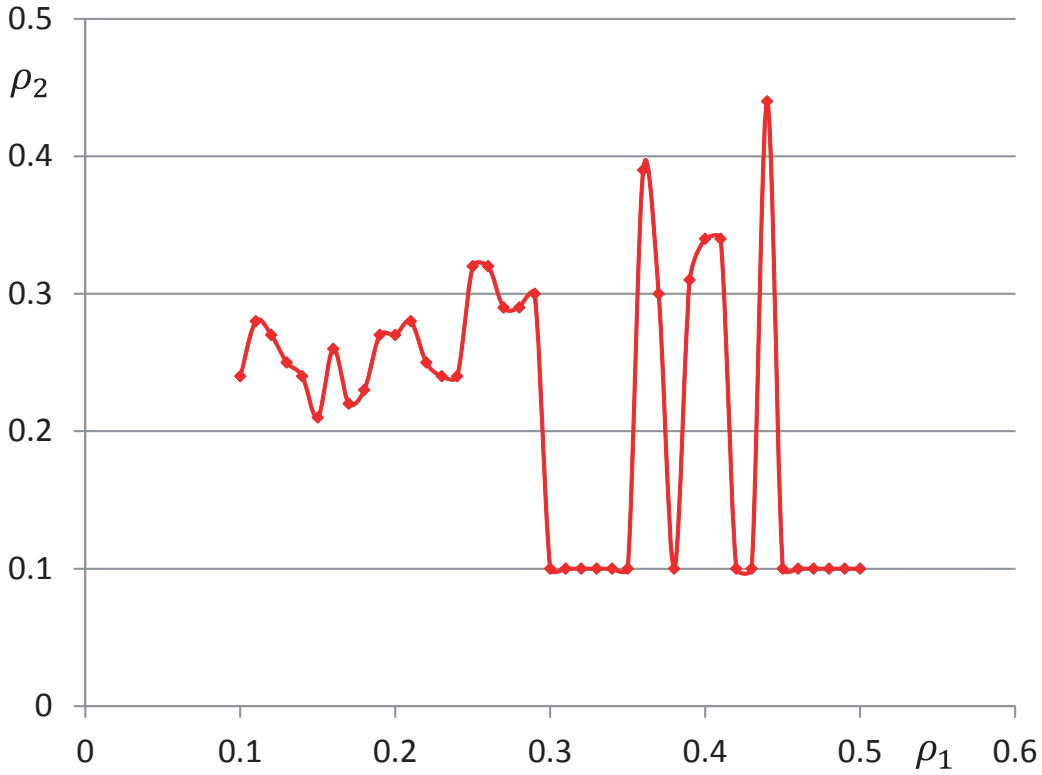


図 3.25: kFCV により決定した ρ_2 (入力 annulus)

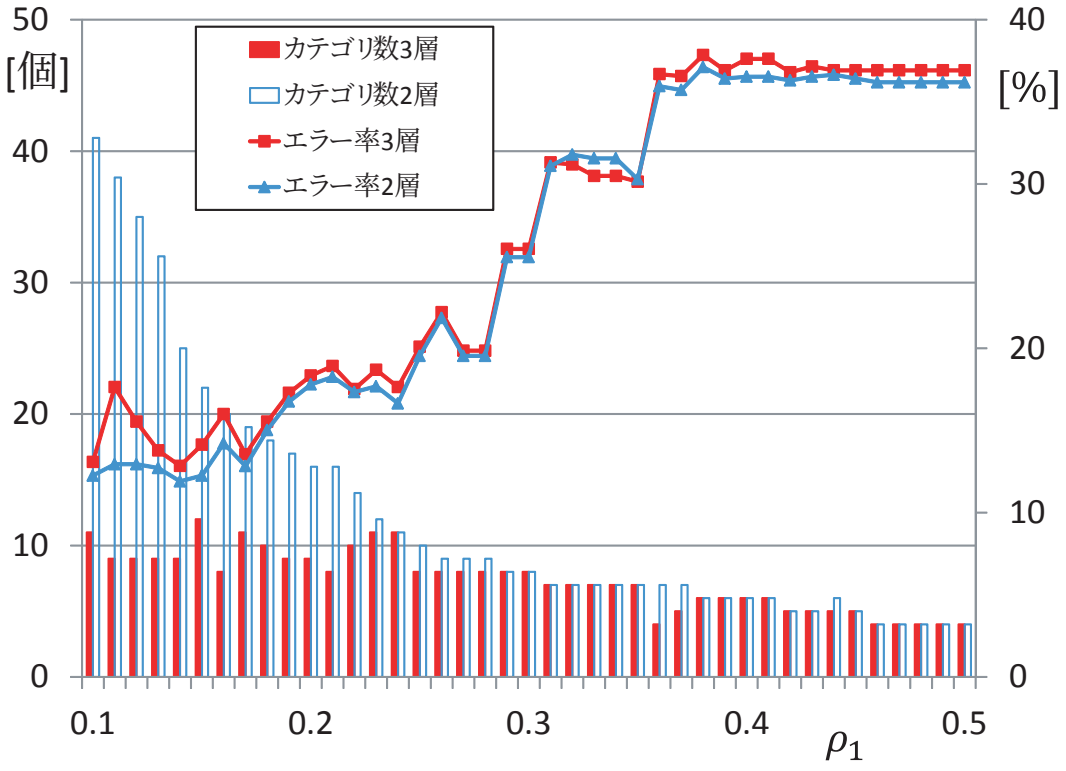


図 3.26: 基本的な RBART-Map と階層型 RBART-Map の分類能力の比較 (入力 annulus)

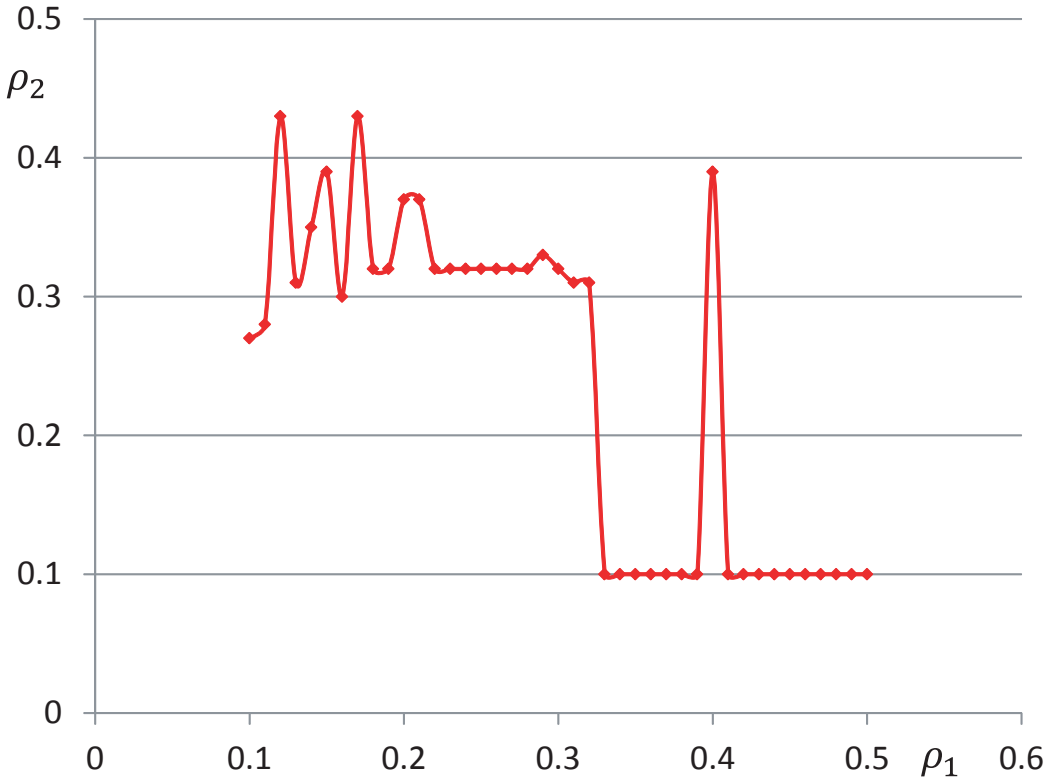


図 3.27: kFCV により決定した ρ_2 (入力 fourclass)

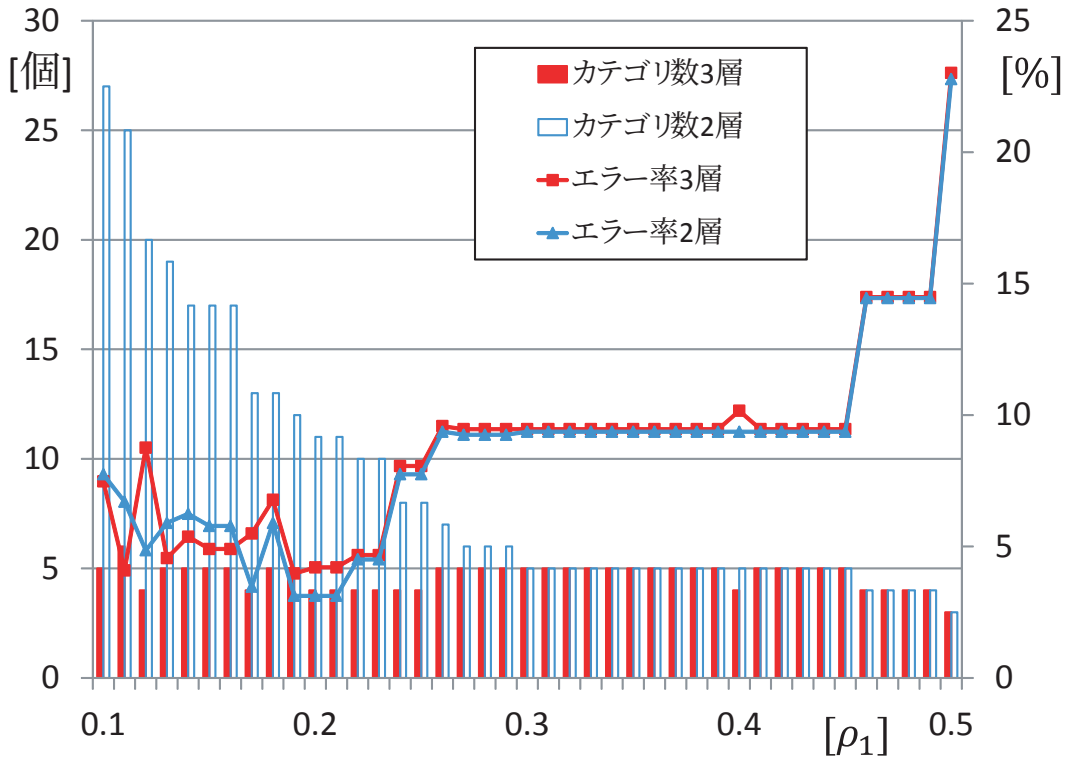


図 3.28: 基本的な RBART-Map と階層型 RBART-Map の分類能力の比較 (入力 fourclass)

第4章　むすび

本論文では，動径基底 ART マップの基本的な分類能力について考察した。特に入力順序が変化することによる、分類能力への影響について検証した。数値実験から、RBART-Map は入力順序が変化することで分類能力に大きな差がでていることがわかる。この結果から、適切な入力順序に変えてから入力を行うことで、より高い分類機能の RBART-Map とすることができる。そして入力順序が変化することでなぜ分類能力が変化するのか考察した。

第 3 章では，階層型 RBART-Map を提案しこのアルゴリズムを示す。基本的な 2 つのベンチマークを用いて基本的な数値実験を行うことで、パラメータ設定による分類能力の解析を行なった。そして基本的な RBART-Map とこれを比較した。その結果以下のことがわかる。

- ビジランスパラメータ ρ_1 が小さいとき階層型 RBART-Map のほうが基本的な RBART-Map よりカテゴリ数を大幅に少なくできる。またこの時エラー率もほぼ変わらない。
- ビジランスパラメータ ρ_1 が大きいいとき階層型 RBART-Map の影響はなく、基本的な RBART-Map と変わらない。
- 交差検証を用いてビジランスパラメータ ρ_2 を自動的に決める手法を提案した。ρ_1 が小さい時、ρ_2 の決定に有効な可能性があることを示した。

今後の課題として、最適な入力順序の決定方法や実問題への応用などがあげらられる。

参考文献

[1] S. Grossberg, “Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions,” Biological Cybernetics, 23, 187-202, 1976.

[2] G.A.Carpenter, S.Grossberg, “A massively parallel architecture for a selforganizing neural pattern recognition machine,” Computer Vision, Graphics, and Image Processing, 37, 54-115, 1987.

[3] G. A. Carpenter, S. Grossberg, D. B. Rosen, “Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system,” Neural Networks, 4, 759-771, 1991.

[4] G. A. Carpenter, S. Grossberg, J. H. Reynolds “ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Oganizing Neural Network,” Neural Networks, 4, 565-588, 1991.

[5] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, D. B. Rosen, “Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps,” IEEE Trans. Neural Networks, 3, 5, 698-713, 1992.

[6] G. C. Anagnostopoulos, M. Georgiopoulos, “Ellipsoid ART and ARTMAP for Incremental Clustering and Classification,” in Proc. IEEE-INNS Int ’l. Joint Conf. Neural Networks, 2, 1221-1226, 2001.

[7] G. C. Anagnostopoulos, M. Georgiopoulos, “Hypersphere ART and ARTMAP for Unsupervised and Supervised, Incremental Learning,” in Proc. IEEE-INNS Int ’l. Joint Conf. Neural Networks, 6, 59-64, 2000.

[8] M. Ohki, H. Torikai and T. Saito, Basic characteristics of a simple radial basis ART network and its application, Proc. of ICCNS (abstract), oral, May 22 (2004).

[9] Y. Enosawa, H. Matsushita, Y. Tanaka, N. Kuji and T. Saito, “Classification Function of Radial-Basis Adaptive Resonance Theory Map and Its Application,” Journal of Signal Processing, 16, 4, pp. 307-310, 2012.

[10] D. Charalampidis, T. Kasparis, and M. Georgiopoulos, “Classification of Noisy Signal Using Fuzzy ARTMAP Neural Networks,” Neural Networks, vol. 12, pp. 1023-1036, 2001.

[11] H. T. Abdelwahab, and M. A. Abdel, “Investigating Driver Injury Severity in Traffic Accident Using Fuzzy ARTMAP,” Journal of Computer-Aided Civil and Infrastructure Engineering, vol. 17, pp. 396-408, 2002.

[12] M. Takanashi, H. Torikai, T. Saito, “An approach to fusion of growing self-organizing maps and adaptive resonance theory maps,” IEICE Trans. Fundamentals, E90-A, 9, 2047-2050, 2007.

[13] Bradley Efron, “Estimating the error rate of a prediction rule: improvement on cross-validation,” Journal of the American Statistical Association, Vol. 78, pp. 316-331, 1983.

[14] 石井健一郎，上田修功，前田英作，村瀬洋, “わかりやすいパターン認識,” オーム社，2009.

[15] UCI Repository of Machine Learning UCI Databases, http://www.ics.uci.edu/ mlearn/MLRepository.html

（国内発表）

岡本祐輔, 江野澤瑶子, 斉藤 利通, “動径基底 ART ネットワークとその応用,” 電子情報通信学会ソサイエティ大会, A-2-13, 北海道, 2011-9

岡本祐輔, 中山雄太, 江野澤瑶子, 斉藤 利通, “ART-Map に基づく論理合成,” 電子情報通信学会技術研究報告, NLP2011-89, 静岡県, 2011-10

岡本祐輔, 斉藤 利通, “動径基底 ART 系の Vehicle Routing Problem への応用について,” 電子情報通信学会総合大会, D-2-6, 岡山県, 2012-3

岡本祐輔, 斉藤 利通, “Radial Basis ARTMAP の改良アルゴリズムの提案,” 電子情報通信学会ソサイエティ大会, a-2-10, 富山県, 2012-9

岡本祐輔, 斉藤 利通, “階層型 ART-Map の分類機能について,” 電子情報通信学会技術研究報告, NC2012-77, 愛知県, 2012-12

謝辞

本修士論文は，著者が法政大学大学院工学研究科電気工学専攻修士課程在学中の 2 年間，同大学理工学部電気電子工学科教授斎藤利通博士の指導の下で行ったものである．研究を遂行するに当たり，終始懇切に御指導，御鞭撻下さった斉藤利通博士に深謝致します。

また、研究活動中に貴重な御助言・御討論を賜りました株式会社東芝 江野澤瑶子氏には感謝の意を表明いたします。

最後に法政大学理工学部電気電子工学科斎藤利通研究室の皆様には著者が在学中，御討論，御助言を頂きました．ここに深謝致します.