

# 属性の重要度を考慮した適応共鳴理論に基づく分類規則の学習

那須 曜 上田 祐彰 高橋 健一 宮原 哲浩  
(広島市立大学大学院 情報科学研究科)

## 1 はじめに

適応共鳴理論[1]に基づいた2つの分類規則学習法 ARTMAP<sub>ED</sub>と ARTMAP<sub>AW</sub>を提案する。

ARTMAP<sub>ED</sub>はユークリッド距離に基づいて分類規則を生成し, ARTMAP<sub>AW</sub>は属性の重要性を分類結果に反映できるように ARTMAP<sub>ED</sub>を改良した手法である。

## 2 ARTMAP<sub>ED</sub>

ARTMAP<sub>ED</sub>は, 2つの ART<sub>ED</sub>をマップフィールドによって結合することによって構築されている(図1)。ARTMAP<sub>ED</sub>の入力は, 属性値ベクトル  $X=(x_1, \dots, x_M)$ およびその分類結果  $R$ によって表現される事例  $I=(X, R)$ である。はじめに,  $X$ が ART<sub>ED</sub>-a および ART<sub>ED</sub>-b に受理され, 各 ART<sub>ED</sub>において  $I$ の属しているカテゴリ  $n_a, n_b$ を求める。なお, カテゴリ  $n_a$ は中心ベクトル  $C_j=(c_{1,j}, \dots, c_{M,j})$ , 半径  $r_a$ の超球によって表現され, 次式を満たすときに  $n_a$ が  $I$ が属しているカテゴリとする。

$$\sqrt{\sum_i (c_{i,j} - x_i)^2} < r_a \quad (1)$$

また,  $r_a$ には初期値として大きな値が設定され,  $r_b$ には小さな値が設定されているとする。次に,  $n_a$ に属する事例の中で出現頻度が最大となっているクラス  $R_a$ ( $n_a$ の最頻クラス)と  $n_b$ の最頻クラス  $R_b$ を比較する。 $R_a=R_b$ のときは,  $R_a$ を  $I$ の推定クラスとして出力し,  $C_j$ を更新する。 $R_a \neq R_b$ のときは,  $I$ が  $n_a$ に属さなくなるまで  $r_a$ を小さくする(マッチトラッキング)。この操作を ART<sub>ED</sub>-a が提示するカテゴリの最頻クラスが  $R_b$ と等しくなるまで実施し,  $R_b$ を  $I$ に対する推定クラスとして出力する。これら一連の操作をすべての事例に対して繰り返し適用することにより, 事例に対する分類規則を学習する。なお, 最終的な分類規則は, ART<sub>ED</sub>-a において各事例を属性値によってカテゴリに分類する分類規則, および各カテゴリの最頻クラスによって表現する。

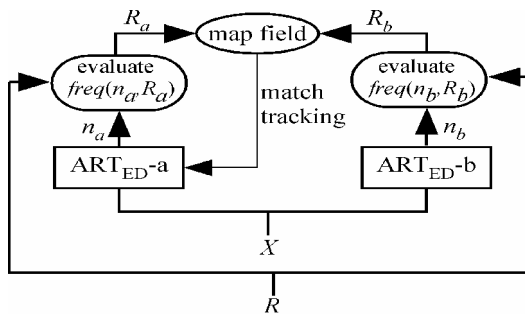


図1: ARTMAP<sub>ED</sub>の構造

## 3 ARTMAP<sub>AW</sub>

一般に, 属性には分類結果に対する影響が大きい属性とそうでない属性がある。しかし ARTMAP<sub>ED</sub>が生成する分類規則(カテゴリに相当)では, 属性の分類結果に対する重要性(以下, 属性の重み)が考慮されていない。そこで, 式(1)の  $1/r_a^2$ を各属性の重み  $w_{i,j}$ に置き換えた式(2)に基づいて

分類規則を学習する ARTMAP<sub>AW</sub>を実装した。 $w_{i,j}$ はマッチトラッキングを行う際および次に述べるカテゴリ統合を行う際に更新している。

$$\sqrt{\sum_i w_{i,j} (c_{i,j} - x_i)^2} < 1 \quad (2)$$

## 4 カテゴリの統合

ARTMAP<sub>ED</sub>および ARTMAP<sub>AW</sub>では, 分類結果に相当するカテゴリが非常に多く生成されるなどの問題が生じる。そこで本研究では, 分類規則の汎化に相当するカテゴリの統合技法[2]を各提案手法に実装した。

カテゴリの統合技法とは, 以下の2つ条件を双方とも満たす2つのカテゴリを統合し, 2つのカテゴリの領域を包含するような最小の超球を統合後のカテゴリとする操作である。

条件1: カテゴリ同士のユークリッド距離が近い。  
条件2: 最頻クラスが一致し, かつ, その次に出現頻度の高いクラスも一致している。

なお, カテゴリの統合はマッチトラッキングと同様に ART<sub>ED</sub>-a のカテゴリに対してのみに適用する。

## 5 実験結果

実験結果を表1に示す。実験には, UCI Machine Learning Databases で公開されている属性数8, クラスの種類3, 事例数4177の *abalone* と属性数10, クラスの種類7, 事例数214の *glass*を用いた。

表1: 実験結果

データ	手法	正答率	ルール数	CPU(ms)
abalone	C4.5	78.9	343	1130
	ARTMAP <sub>ED</sub>	74.1	190	1763
	ARTMAP <sub>AW</sub>	91.2	145	2123
glass	C4.5	88.9	14	30
	ARTMAP <sub>ED</sub>	72.9	18	26
	ARTMAP <sub>AW</sub>	90.1	12	31

ARTMAP<sub>AW</sub>は, 他の手法に比べて実行時間はかかるが, ルール数が少なく正答率の高い分類規則を生成することが出来た。また, ランダム値をとる属性を追加したデータに対する実験の結果, ARTMAP<sub>AW</sub>では属性の重要度を考慮して, 分類規則を生成しているということが確認できた。

## 6 おわりに

適応共鳴理論を応用した2つの新しい分類規則学習法を提案し, 実験結果を提示した。なお, 本研究は広島市立大学特定研究費(No.7115)による支援を受けた。

## 参考文献

- [1] G.A.Carpenter et.al, "A neural network architecture for incremental supervised learning of analog multi-dimensional map," *IEEE Trans. on Neural Networks*, Vol.3, No.5, pp.698-713, 1992.
- [2] H.Ueda et.al, "Categorization of continuous numeric percepts by modified fuzzy ART with Q-learning," *Int. J. Pattern Recognition and Artificial Intelligence*, Vol.20, No.2, pp. 113-127, 2006.