

4E-12 Conjunctive Network による概念学習

塚本 義明 西村 和幸 生天目 章

防衛大学校 情報工学科

1. はじめに

多層ネットワークによるコネクショニスト学習は、望ましい入出力パターンとの関係について正しく学習する過程において、入出力パターンとの関係について特徴付ける高次の関係式がネットワークの内部表現として獲得される。したがってコネクショニスト学習は、その並列処理や汎化能力の特徴を生かした並列的な概念の学習法として有力な手段である。しかし、概念学習が対象とする領域における知識の対象要素は複雑な構造を有している場合が多く、構造的な知識表現された対象領域におけるコネクショニスト学習の有効性について十分に解明されていない。特に、多層ネットワークの学習アルゴリズムとして逆伝搬法を用いたコネクショニスト概念学習では、概念の複雑な構造関係について正しく推論できないことが示されている。(1)(2)

本論では、コネクショニストモデルによる概念の学習を予備学習段階とネットワーク学習段階の2段階に分けて行なう方法について提案する。予備学習段階とは、与えられた学習例の集合を用いて、対象とする領域の概念学習に適切なネットワークアーキテクチャ(特に隠れユニットの数と構造)について決定するための事前学習フェーズである。

ネットワークの構造について決定するためのアルゴリズムとして、最小連言探索アルゴリズムについて提案する。ネットワーク学習段階では、事前に提案された構造を持つ多層ネットワークを用いて、学習例に示された入出力関係を表現するためのネットワークの結合係数が決定される。

2. 構造のある世界でのコネクショニスト概念学習モデル

それぞれの対象要素(オブジェクト)が n 個の属性変数 $A_i, i=1, 2, \dots, n$, により特徴づけられる世界 W を考える。属性変数は、多値属性で離散値をとる。世界 W での各々のオブジェクトは、ビットベクトルで表わすことができる。各オブジェクトの i 番目の属性は長さ m_i のベクトル α_i で表わされ、ベクトル α_i の j 番目は 1 か 0 を取り、属性の有無を表現する。各オブジェクトは、いくつかのカテゴリ $C = \{C_1, C_2, \dots, C_k\}$

にクラス分けでき、各要素をカテゴリ値と呼ぶ。特徴ベクトル $X = A_1 * \dots * A_n$ の部分集合を S とする。 C の部分集合 $C(S)$ を以下のように定義する: 各要素は S の属性変数からなる事例のカテゴリ値である。例えば、 $S = \{a_1\}$, $a_1 \in A_1$ ならば、 $C(S)$ は属性変数 a_1 をとる全ての例のカテゴリ値の集合である。

学習例の集合 Y は、 $\langle X, C \rangle$ の集合である。ここで、 X と C は、オブジェクトの属性変数とカテゴリ値を表わす。ブール変数の集合である p 番目の学習例を特に、 $\langle X_p, C_p \rangle = \{x_{1p}, x_{2p}, \dots, x_{np}, c_{1p}, \dots, c_{mp}\}$ で表わす。

3. Conjunctive Network

3.1 アーキテクチャ

このネットワークは、3層ネットワークである。入力層は、属性変数に対応させる。隠れ層には、学習例から属性変数の最小連言表現を見つけそれに対応させる。ユニットの数は、発見された連言表現による。入力層と隠れ層間および隠れ層と出力層の結合係数は、ネットワークの学習アルゴリズムにより決定する。

3.2 最小連言表現探索アルゴリズム

C_j の最小連言表現とは、特徴ベクトル X の部分集合 S に対して、

$$\sum_{p: X_i \in S} \prod X_{ip} = \sum_{p: X_i \in S} \prod X_{ip} C_{jp} \dots \textcircled{1}$$

を満たすことである。

連言表現は、 n 個の属性があれば 2^n 個ある。無駄な探索を避けるには、 $\textcircled{1}$ を用いれば効率的に最小連言表現が求まる。

属性の組み合わせ木構造を考える。 $\textcircled{1}$ の条件で1方向横型探索を行ない、探索された連言表現が各ルート属性の最小連言表現集合に属していなければその集合にに入れる。最終的には各ルート属性の最小連言表現集合の和が求める最小連言表現の集合である。

Concept learning with Conjunctive Network

Yoshiaki TSUKAMOTO, Kazuyuki MISHIMURA, Akira NAMATAME

National Defence Academy

この探索法は以下の利点がある。

- 1 探索された組合わせ木の子連言表現をカット可能。
 - 2 1方向により隣接木の連言表現もカット可能。
- ### 3.3 連言層の構築

上記の最小連言表現を中間ユニットとして配置し、連言層と称する。連言層の出力は、ブール値を取り各入学習例を入力とした各連言出力となる。もし最小連言表現発見アルゴリズムがK個の連言を発見したならば、連言層にK個のユニットを配置する。

4. 検証

4.1 具体例

学習対象関数として、 $C = X_1X_2 + X_3X_4 + X_5$ で表現されるブール関数を考える。5変数は32個のパターンを持つがその中から任意に24個及び16個を学習例として抽出し、Conjunctive Networkによる学習能力、汎化能力について評価した。

(1) ケース1：(学習例が24個)

正の学習例を特徴づけるために最小連言探索アルゴリズムが求めた連言形式は、

$$X_5, X_1X_2, X_2X_3, X_2X_4, X_3X_4$$

負の学習例を特徴づけるために求めた連言形式は

$$X_2X_5, X_1X_3X_5, X_1X_4X_5, X_3X_4X_5$$

である。

(2) ケース2：(学習例が16個)

正の学習例を特徴づけるために最小連言探索アルゴリズムが求めた連言形式は、

$$X_5, X_1X_2, X_2X_3, X_2X_4, X_3X_4, X_1X_2X_3$$

負の学習例を特徴づけるために求めた連言形式は、

$$X_2X_5, X_3X_5, X_1X_3X_5, X_1X_4X_5$$

である。

4.2 検証結果

全ての学習例が提示されていないため、学習例にないパターンについて全て正しく分類していない。特に、正の学習例及び負の学習例についてOver-generalizationをする。しかし、中間ユニット及び出力ユニット内でOver-generalizationに対して相殺する効果がある。

ネットワーク規模を同じにした逆伝搬法のネットワークと比較したものが表1である。

構造のある領域では、従来のアーキテクチャより構造を考えた今回提案したConjunctive Networkが有利であることがわかる。

5. 結び

我々は、構造のある領域でのコネクショニスト学習を研究してきた。構造のある領域では、今までの学習アーキテクチャでは不相当であることを示し、新しいアーキテクチャであるConjunctive Networkを提案した。この新しい学習アルゴリズムは、入力属性から最小連言表現を発見することが特徴である。

我々は、今回小さい概念で検証し、伝統的なアルゴリズムより汎化能力が高いことが明らかになった。

学習例	テストデータ	BP		CN	
		誤り	総数	誤り	総数
16	16	2	1	0	2
24	8	4	1	0	2

表1 逆伝搬法とConjunctive Networkの誤り比較

参考文献

- 1) Namatame, "Structured Neural Networks and Thier FlashLearning", Proc. of ALT '90 p67-80
- 2) S.E.Hamppson and D.J.Volper, "Representing and Learning Boolean Function of Multivalued Features" IEEE Transactions on Systems, Man, and Cybernetics, Vol.20, No.1 '90 p67-79