

종족 유전 알고리즘을 이용한 MLP 분류기의 구조학습

(A Structural Learning of MLP Classifiers Using Species Genetic Algorithms)

愼晟孝*, 金商雲*

(Seong-Hyo Shin and Sang-Woon Kim)

요 약

주어진 응용에 적합한 MLP 분류기의 위상 구조를 학습하기 위하여 유전 알고리즘을 이용하는 방법이 있다. 그러나 이 방법은 고차원 다클래스의 실제 응용에 적용할 경우 탐색 공간이 특징 차원의 수에 대해 지수 함수적으로 급증하게 된다. 따라서 방대한 탐색 공간을 효율적으로 학습하기 위하여 종족 유전 알고리즘(SGA)을 이용한 학습 방법을 제안한다. SGA에서는 전체 탐색 공간을 은닉층 뉴런의 수에 따라 여러 개의 영역으로 분할하여, 이를 “종족”이라 한다. 그리고 학습 과정 중 기준에 미달되는 종족에 대해서는 학습을 중단시킴으로써 불필요한 탐색을 줄인다. 실험 결과 제안한 방법은 기존의 방법에 비해 오인식율, 학습 속도, 위상 구조 등에 있어서 효율적임을 확인하였다.

Abstract

Structural learning methods of MLP classifiers for a given application using genetic algorithms have been studied. In the methods, however, the search space for an optimal structure is increased exponentially for the physical application of high dimension-multi class. In this paper we propose a method of MLP classifiers using species genetic algorithm(SGA), a modified GA. In SGA, total search space is divided into several subspaces according to the number of hidden units. Each of the subdivided spaces is called “species”. We eliminate low promising species from the evolutionary process in order to reduce the search space. Experimental results show that the proposed method is more efficient than the conventional genetic algorithm methods in the aspect of the misclassification ratio, the learning rate, and the structure.

I. 서 론

다층신경망(Multi-Layered Perceptron; MLP) 분류기의 학습에 영향을 미치는 요소는 MLP의 위상 구조, 학습 계수, 활성화 함수, 관성 계수, 초기 가중치 등 여러 가지가 있다. 이 중에서 위상 구조는 학습에 가장 큰 영향을 미치는 요소로서, 입력층, 은닉층, 출력층 뉴런의 수를 각각 I , H , O 라고 할 때, I 는 특징

차원의 수에 따라 결정되며 O 는 클래스의 수로 결정된다. 그러나, 주어진 응용에 적합한 H 는 설계자의 경험에 의하여 결정되고 있다.

이론적으로는 선형분리 가능한 클래스의 수를 k 라 할 때 최소 뉴런의 수는 $H_{\min} = \lfloor \log_2 k \rfloor$ 이고 최대 $H_{\max} = k - 1$ 이다^[1]. 실제로 주어진 응용에 적합한 최적 구조를 학습하는 방법은 최소한의 구조로부터 시작하여 점차 성장(growing)시켜가는 방법^[2]과 충분히 큰 구조로부터 점차 삭제(pruning)하여 가는 방법^[3]이 있다. 그러나, 이러한 순차적인 방법에서는 초기 구조의 설정에 따라서 최적구조에 수렴하지 않거나 지역

* 正會員, 明知大學校 컴퓨터工學科

(Dept. of Computer Eng., Myongji-Ji Univ.)

接受日 1996年8月14日, 수정완료일: 1998年1月30日

적 최소값에 빠지는 등의 문제가 있으며, 삭제 또는 성장의 초기 단계에 불필요한 연산을 많이 수행하게 된다.

이러한 문제를 해결하기 위하여 유전 알고리즘(Genetic Algorithm; GA)을 이용하는 방법이 연구되었다. 처음에는 GA를 이용하여 MLP의 가중치만을 구하는 연구였으나^[4], 점차 위상 문자열과 가중치 문자열을 이용하여 MLP의 구조를 자동적으로 결정하는 연구^[5]로 발전하였다. 그러나 이 방법으로 도출한 MLP는 필요이상의 은닉층 뉴런을 갖고 있으며, 또한 완전 연결 구조(fully connected structure)로써 불필요한 링크를 포함한다. 따라서, 꼭 필요한 링크들만을 갖는 선택적 연결(selectively connected structure)의 MLP를 구하는 연구가 이루어졌다^[5]. 그런데 이 방법에서는 가중치뿐만 아니라 링크의 필요여부까지 학습해야 하기 때문에 많은 학습 시간을 요구한다.

따라서, 이 논문에서는 GA에 종족(species) 개념^[6]을 도입하여 보다 빠른 시간 내에 최적 구조를 도출할 수 있는 종족 유전 알고리즘(Species Genetic Algorithm; SGA) 방법을 제안한다^[7]. SGA는 전체 탐색공간을 은닉층 뉴런의 수에 따라 종족으로 분할한 다음, 불필요하다고 판단되는 종족에 대해서는 탐색을 중단함으로써 전체적인 학습 시간을 줄이는 방법이다. 이 때, 종족의 필요성 여부는 학습율을 보고서 결정한다.

한편, Srinivas^[8] 등은 MLP의 구조적 특성을 이용하여 GA의 탐색 공간을 축소하는 방법을 제안하였다. 즉, 은닉층 뉴런의 입력력 링크의 가중치와 다른 은닉층 뉴런의 입력력 링크의 가중치가 서로 바뀌어도 출력값은 같게 되는 대칭성을 이용하여 중복 탐색을 방지하는 방법이다. 따라서 이 연구에서는 제안 방법을 평가하기 위하여 단순 GA와 Srinivas법을 결합한 방법 및 SGA와 결합한 방법과 성능을 비교한다.

이하 제 2장에서는 GA 및 SGA를 이용하여 MLP 분류기의 구조를 학습하는 방법에 대하여 설명한다. 그리고 제 3장에서는 실제 데이터에 대한 실험 결과를 분석 평가하고, 마지막으로 제 4장에서 결론을 맺는다.

II. SGA 학습 방법

1. GA를 이용한 구조 학습 방법

최적 구조의 MLP 분류기를 도출하기 위한 GA의

염색체 구조는 그림 1과 같다. 여기서, 위상 문자열(topology string)은 은닉층 뉴런의 수 H 를 나타내고, 태그 문자열(tag string)은 각 링크의 필요 여부를 판단하기 위한 것이다. 해당 순서의 비트가 '1'이면 연결 링크이고 '0'이면 연결되지 않은 링크를 나타낸다. 가중치 문자열은 각 링크의 가중치를 표현하며 실수로 이루어져 있다.

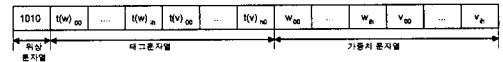


그림 1. GA에서의 염색체 구조

Fig. 1. The representation of chromosomes in GA.

GA를 이용하는 방법에서는 그림 1의 염색체 집단을 생성, 교배, 돌연변이의 유전 연산자를 적용하여 진화시킴으로써 최적구조를 학습하게 된다. 그러나, GA의 탐색 공간은 위상 구조의 크기에 지수함수적으로 증가하기 때문에, 고차원 응용의 경우에는 탐색 공간을 줄여 학습 시간을 단축해야만 한다.

2. SGA를 이용한 학습 방법

Potts 등^[6]은 GA의 학습 속도 및 최적해 탐색에 대한 성공율을 높이기 위하여 종족(species)개념을 도입하였다. 문헌 [6]에서는 전체 염색체를 세 개의 종족으로 나누어 학습을 수행하며, 각 종족마다 교배, 돌연변이 확률 등의 GA 파라메타를 달리하였다. 즉, 여러 가지 환경에서 염색체들이 진화할 수 있게 함으로써 기존의 방법들의 문제점을 개선하였다.

따라서 이 논문에서는 종족 개념을 이용한 유전 알고리즘 SGA로 위상 구조를 학습하는 방법을 제안한다. SGA는 위상 구조의 탐색 공간을 은닉층 뉴런의 수 H 를 기준으로 그림 2와 같이 여러 개의 부분 집합으로 나누고, 이를 종족이라 명명한다. 여기서, 종족 S_H 는 동일한 H 를 갖는 염색체들의 집합으로, H 개의 은닉층 뉴런 구조의 가중치 및 링크의 필요 여부 등을 학습하게 된다. S_{-1} 은 각 종족에서의 최고 염색체만을 선택하여 보관하는 종족으로써 이후 소개될 재시도 연산에 이용된다. H_{min} 은 최소 은닉층 뉴런의 수이고, H_{max} 는 최대 은닉층 뉴런의 수이다. 또한 세 종류의 선으로 나타낸 이주, 선택, 재시도 연산은 SGA에서 새롭게 정의된 연산자로서 다음과 같다.

(1) 이주 연산(migration) : 한 종족내의 염색체를 다른 종족으로 옮기는 연산이다. 이주 연산의 목적은

학습이 잘 이루어지는 종족으로 염색체를 이주시키는 데 있다. 따라서, 학습이 잘 이루어지는 종족일수록 염색체를 많이 보유할 수 있으며, 또한 학습 속도를 향상시킬 수 있다.

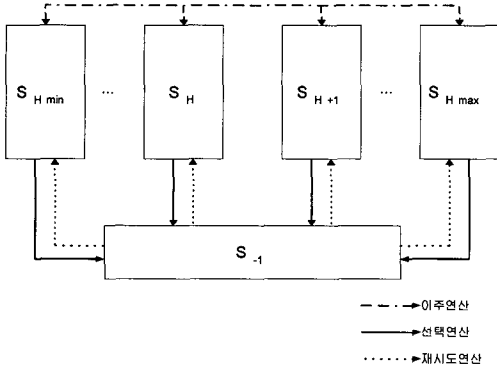


그림 2. SGA로 학습하는 전체 구조
Fig. 2. The whole learning structure in SGA.

어떤 종족에 이주 연산을 적용할 것인지를 판단하기 위하여 식 (1)의 이주 확률을 사용한다. 이주 확률은 교배, 돌연변이 확률처럼 $S_{H_{min}}$ 과 $S_{H_{max}}$ 사이에서 랜덤하게 염색체에 적용한다.

$$M_H(t) = 1 - a \frac{R_H(t)}{\sum_{i=H_{min}}^{H_{max}} R_i(t)} \tag{1}$$

여기서, t : 진화세대,

$M_H(t)$: t 세대에서 S_H 의 이주 확률,

$R_H(t)$: t 세대에서 S_H 의 학습율,

a : 상수.

학습율은 각 종족이 학습한 양을 나타내는 척도로써 식 (2)와 같이 구한다.

$$R_H(t) = \max(F_H(t^*)) - \min(F_H(t^{**})), \quad (0 \leq t^*, t^{**} \leq t; H_{min} \leq H \leq H_{max}) \tag{2}$$

여기서, $R_H(t)$: t 세대에서 S_H 의 학습율,

$F_H(t)$: t 세대에서 S_H 의 최고 염색체의 적합도 값.

(2) 선택 연산(selection) : 각 종족에서 최고 적합도를 갖는 염색체를 선택하는 연산이며, 선택된 염색체들은 S_{-1} 에 보관한다. 여기서, 염색체의 적합도(fitness) 식 (3)과 같이 최소 자승 오차(mean squared error ; MSE)와 엔트로피(entropy)를 이용

하여 구한다.

$$F_H^k(t) = \frac{1}{MSE_H^k(t) + \lambda \times ETRP_H^k(t)} \tag{3}$$

여기서, t : 진화 세대,

k : 염색체 번호,

$\lambda = \lambda_0 R_H(t)$, λ_0 : 초기 상수,

$F_H^k(t)$: t 세대에서 종족 S_H 의 k 번째 염색체의 적합도,

$MSE_H^k(t)$: t 세대에서 종족 S_H 의 k 번째 염색체의 MSE,

$ETRP_H^k(t)$: t 세대에서 종족 S_H 의 k 번째 염색체의 엔트로피.

엔트로피는 MLP의 복잡도를 나타내는 척도^[9]로써, 식 (4)와 같이 구한다.

$$ETRP_H^k(t) = \sum_{i,j} P_{ij} \log_2 P_{ij}, \tag{4}$$

$$P_{ij} = \frac{|w_{ij}|}{\sum_{i,j} |w_{ij}|}.$$

여기서, w_{ij} : i 번째 뉴런과 j 번째 뉴런사이의 가중치.

(3) 재시도 연산(reintroduction) : 일정 기간 학습이 제대로 이루어지지 않을 경우 기존의 염색체들을 다시 초기화하여 재학습하는 연산이다. 이 때, S_{-1} 에 있는 최고 염색체들을 바탕으로 초기화함으로써 그동안 학습했던 정보들을 유지할 수 있다. $H_{min}(t+1) = H_{max}(t) + 1$, $H_{max}(t+1) = H_{max}(t) + (H_{max}(t) - H_{min}(t))$ 로 설정하여 초기의 학습 때보다 크게 한 후, S_{-1} 내 염색체들의 스트링에 작은 수의 랜덤값을 더하여 초기화시킨다.

SGA에서의 염색체 표현은 그림 1의 GA 염색체와 유사하나 종족 자체가 은닉층 뉴런의 수를 표현하고 있기 때문에 위상 문자열은 필요가 없다. 따라서, 염색체는 그림 3과 같이 정의한다. 그림에서 염색체의 길이는 은닉층 뉴런의 수 H 에 의하여 결정된다.

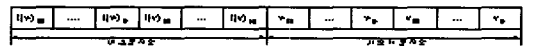


그림 3. SGA에서의 염색체의 표현
Fig. 3. The representation of chromosomes in SGA.

3. 불필요한 종족의 제거

그림 2에서 초기에 임의로 설정한 H_{min} 과 H_{max} 사

이의 모든 위상 구조를 학습하는 것은 시간을 많이 요하는 작업이다. 따라서, 학습하는 도중에 학습이 불필요하다고 판단되는 종족에 대해서는 학습을 중지시킨다. 이 때, 불필요한 종족이란 MLP의 위상 구조가 최적의 구조보다 작거나 크다고 판단되는 종족으로 아래의 두 가지 규칙에 의해 판별한다.

우선 최적보다 작은 종족의 판단은 [규칙 1]에 의한다. 학습이 가장 부진한 경우는 H 가 작아서 입출력 관계를 실현하지 못할 때이며, H 가 작을수록 학습율은 더 떨어지게 된다. [규칙 1]은 전체 종족 중 학습율이 가장 안 좋은 종족 S_l 을 찾아내는 규칙이며, S_l 이하의 종족을 불필요한 종족으로 판단한다.

[규칙 1] $S_l = \arg \min_H (R_H(t)); S_{H_{min}} \sim S_l$ 사이의 종족을 최적의 H 보다 작다고 판단한다;

최적보다 큰 종족의 판단은 [규칙 2]에 의한다. 한 종족 S_m 의 학습율이 H 가 큰 종족 S_n 의 학습율보다 좋다면, 종족 S_m 은 종족 S_n 보다 최적의 구조에 인접하다. 따라서, 은닉층 뉴런의 수가 n 개 이상인 구조에 대해서는 학습할 필요가 없으며 S_n 이후의 종족을 학습에 불필요한 종족으로 판단한다.

[규칙 2] $if(R_m(t) \geq R_n(t), m < n)$ then $S_n \sim S_{H_{max}}$ 사이의 종족을 최적의 H 보다 크다고 판단한다;

위의 두 규칙에 의해 불필요하다고 판단되는 종족들은 제거되며, [규칙 1]에 의해 $H_{min} = l+1$, [규칙 2]에 의해 $H_{max} = n-1$ 이 된다. 이 규칙들은 매 세대 적용하는 것이 아니라 규칙 적용 세대(rule-apply-generation)에만 적용한다. 규칙 적용 세대는 다음과 같이 결정한다.

$if(\max_H(R_H(t)) \geq \max_H(R_H(t^{***})) \times 2)$ then t is rule-apply-generation;

즉, t 세대에서 가장 좋은 종족의 학습율이 이전에 규칙을 적용했던 세대(t^{***})의 학습율보다 2배 이상 증가하면 t 를 규칙 적용 세대로 한다.

SGA의 전체 학습 과정은 그림 4와 같다. 여기서

α, β, γ 는 설계자가 정의하는 실험 상수이고, t, tt 는 세대수이다. 단계 1에서 14까지의 내용은 GA와 유사하다. 단, 종족을 사용하기 때문에 각 종족별로 연산을 수행하고, 단계 12에서는 이주 연산이 추가되었다. 단계 15부터 18까지는 SGA에서 새롭게 정의하여 사용되는 연산이며, 앞에서 설명한 불필요한 종족의 제거 및 재지도 연산이 수행된다.

```

1. SGA algorithm(
2.     t = tt = 0;
3.     H_min = alpha; H_max = beta;
4.     for(H=H_min; H < H_max; H++)(
5.         initialize S_H(t);
6.         evaluate S_H(t);
7.     )
8.     while( not termination_condition )(
9.         t = t + 1;
10.        for(H=H_min; H < H_max; H++)(
11.            select and reproduce S_H(t) from S_H(t-1);
12.            alter S_H(t) with crossover, mutation and migration;
13.            evaluate S_H(t);
14.        )
15.        if( t == rule-apply-generation )
16.            apply [rule 1] and [rule 2];
17.        if(max(F_H(t)) == max(F_H(t-1))) tt = tt + 1;
18.        else tt = 0;
19.        if( tt == gamma )(
20.            alter S_H(t) with reintroduction;
21.            tt = 0;
22.        )
23.    )
24. )
    
```

그림 4. SGA의 학습 과정
Fig. 4. The learning procedure of SGA.

III. 실험 및 결과 고찰

1. 실험 데이터 및 실험 방법

제안 방법의 성능 평가를 위하여 사용한 실험 데이터는 표 1과 같다. 여기서 E13B는 수표에서 사용하는 14가지의 비트맵 문자에 대한 이진패턴이며, LET-TER는 여러 폰트의 알파벳에서 16개의 특징을 추출한 실수패턴이다. 각각의 학습 데이터에 대하여 전체 패턴 중 일부는 학습에 사용하고, 나머지는 테스트에

이용하였다.

표 1. 실험 데이터의 특성

Table 1. The characteristics of three experimental data.

특성 데이터	특징의 수	클래스의 수	패턴의 수		
			전체	학습	테스트
IRIS	4	3	150	75	75
E13B	63	14	5,614	560	5,054
LETTER	16	26	5,200	1,040	4,160

실험에 사용된 파라메타 값들은 다음과 같다.

- 학습 세대=1,000(IRIS), 2,000(E13B, LETTER),
- 돌연변이 확률 $m = 0.01$; 교배 확률 $c = 0.7$; $\lambda_0 = 10^{-3}$,
- $H_{min} = 1, H_{max} = 20, \alpha = 3$,
- 전체 염색체 수 = 200,
- 한 종족의 초기 염색체 수 = 20,
- 한 종족의 최대 염색체 수 = 100.

Sun sparc/20에서 C언어를 이용하였으며, 실험 결과는 세 가지 데이터에 대해 20번씩 수행한 후 평균치를 구하여 표현하였다.

실험 방법은 (1) 단순 GA를 이용하여 최적 구조를 학습하는 방법(GA), (2) 단순 GA와 Srinivas법을 결합한 방법(RGA), (3) SGA를 이용하여 최적 구조를 학습하는 방법(SGA), (4) SGA와 Srinivas법을 결합한 방법(RSGA)으로 하였다.

2. 각 방법의 학습 속도 고찰

우선 각 방법의 학습 속도를 평가하기 위하여 GA, RGA, SGA, RSGA 방법을 같은 시간동안 학습한 후의 MSE를 표 2와 같이 측정하였다. 학습이 빨리 이루어지는 방법일수록 많은 오차를 줄일 수 있기 때문에 MSE가 작게 된다.

모든 데이터에 대하여 GA가 가장 큰 MSE를 보여 학습 속도가 가장 느리며, RSGA가 가장 낮은 MSE를 보여 다른 방법들에 비하여 학습이 잘 이루어진다. 그리고, 표준 편차를 비교해보면 SGA와 RSGA가 GA와 RGA에 비해 낮은 것을 볼 수 있다.

대칭성을 이용하는 방법인 RGA와 RSGA가 대칭성을 이용하지 않는 GA와 SGA의 성능에 비해 우월하다. 따라서, 중복 탐색을 피할 수 있는 방법이 효과

를 보인다.

표 2. MSE 비교

Table 2. Comparisons of the MSE.

데이터 방법	IRIS	E13B	LETTER
GA	0.026 ± 0.023	0.054 ± 0.094	0.104 ± 0.143
RGA	0.023 ± 0.021	0.047 ± 0.084	0.098 ± 0.104
SGA	0.009 ± 0.012	0.019 ± 0.041	0.063 ± 0.076
RSGA	0.008 ± 0.010	0.015 ± 0.041	0.057 ± 0.069

종족을 이용하지 않는 GA와 RGA에 비해 종족을 이용한 SGA와 RSGA의 MSE가 훨씬 낮다. 따라서, 종족에 의한 불필요한 영역의 제거방법이 학습 속도에 많은 영향을 미친다. 이 과정을 보이기 위하여 그림 5와 같이 IRIS에 대하여 GA, SGA로 학습하였을 경우 염색체의 종류가 수렴하는 정도를 나타내었다. 여기서 염색체 종류는 은닉층 뉴런의 수에 따라 결정되는 것으로서, 염색체의 종류가 수렴하였다는 것은 은닉층의 뉴런 수가 결정되었음을 의미한다. SGA의 경우 불필요한 종족의 제거 규칙으로 인하여 GA보다 안정되게 수렴함을 보여준다.

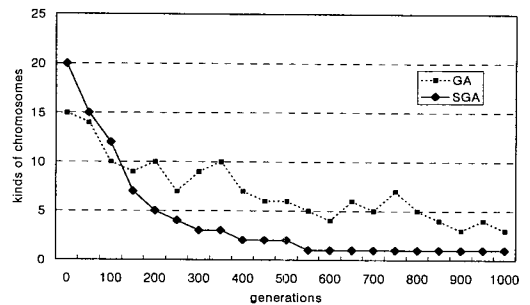


그림 5. 염색체 종류의 수렴율 비교

Fig. 5. Comparisons of chromosome-convergence-rate for IRIS.

표 3(a), (b)는 IRIS에 대하여 GA, SGA로 학습할 때 각 뉴런수의 염색체가 변하여 가는 과정을 보여준다. 예를 들어 표 3(a)는 GA로 학습할 때 초기에 뉴런 수가 1, 2, ..., 10인 염색체를 각각 10개씩 설정한 후 100세대 진화시키면 각 뉴런 수의 염색체가 0, 18, 25, 23, 19, 21, 14, 8, 7, 9개로 변하고, 900세대 후에는 2, 3, 4 뉴런의 염색체만 56, 51, 43개 남고 나머지 뉴런수의 염색체는 도태함을 나타낸다.

또한, 표 3(b)는 SGA법의 경우 600세대 진화 후 은닉뉴런이 세 개인 염색체만 100개 생성함을 나타낸다. 즉 표 3은 IRIS의 경우 SGA법이 GA에 비해 빨리 은닉층 뉴런이 세 개인 위상구조를 찾아냄을 나타낸다.

표 3. 유닛의 수 수렴 비교

Table 3. Comparisons of units-convergence for IRIS.

세대 뉴런의 수	세대								
	100	200	300	400	500	600	700	800	900
1	0	0	3	0	0	0	4	4	0
2	18	24	25	33	39	39	42	45	56
3	25	34	31	41	43	53	57	55	51
4	23	31	35	28	33	31	29	25	43
5	19	24	21	21	18	27	18	21	0
6	21	12	15	12	13	0	0	0	0
7	14	9	11	10	4	0	0	0	0
8	8	7	6	5	0	0	0	0	0
9	7	3	3	0	0	0	0	0	0
10	9	4	0	0	0	0	0	0	0

(a) GA

세대 뉴런의 수	세대								
	100	200	300	400	500	600	700	800	900
1	0	0	0	0	0	0	0	0	0
2	35	49	53	100	100	0	0	0	0
3	36	78	100	100	100	100	100	100	100
4	34	36	47	0	0	0	0	0	0
5	15	27	0	0	0	0	0	0	0
6	16	10	0	0	0	0	0	0	0
7	12	0	0	0	0	0	0	0	0
8	11	0	0	0	0	0	0	0	0
9	9	0	0	0	0	0	0	0	0
10	10	0	0	0	0	0	0	0	0

(b) SGA

3. MLP 분류기 구조 고찰

실험 데이터에 대하여 각각의 방법으로 학습한 MLP분류기의 구조는 표 4와 같다. 각 방법에 있어서 은닉층 뉴런의 수는 별 차이가 없으나, 전체 링크의 수에 있어서는 차이를 보인다. 전체 링크의 수에 대한 차이는 3.2절의 MSE와 마찬가지로 대칭성을 이용한 탐색 공간 축소 방법을 사용하거나 종족을 이용하는 방법이 보다 효율적이다.

표 4. MLP분류기의 구조 비교

Table 4. Comparison of the structure of MLP classifiers.

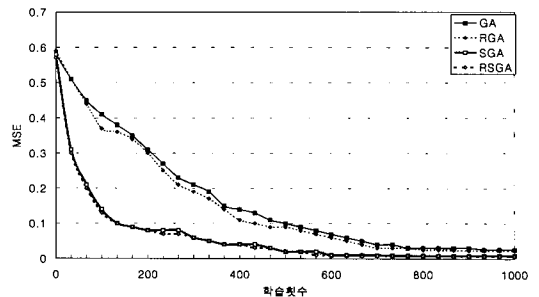
구조 방법	IRIS		E13B		LETTER	
	H	L	H	L	H	L
GA	2.56	12.44	7.01	278.42	12.31	254.21
RGA	2.62	12.62	6.97	264.84	12.21	247.66
SGA	2.63	12.18	6.58	208.91	13.01	209.57
RSGA	2.51	12.03	6.61	207.55	12.68	205.41

(H : 은닉층 뉴런의 수, L : 전체 링크의 수)

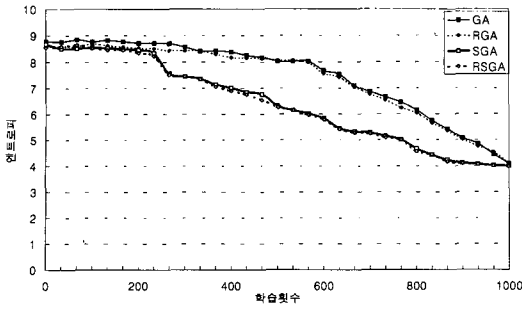
MSE에서 우수한 성능을 가지는 방법일수록 전체 링크의 수 또한 적은 것을 표 2와 표 3을 통해서 알 수 있다. 구조 학습과 MSE의 관계를 나타내기 위하여 IRIS 데이터에 대한 MSE와 엔트로피 값의 감소 과정을 그림 6과 같이 구하였다.

모든 방법이 학습 초기에는 엔트로피의 감소율이 정지상태를 보여 구조를 줄이는 학습은 일어나지 않는 반면, MSE를 줄이는 학습이 주로 이루어진다. 이와는 반대로 학습 후반에는 MSE를 줄이는 학습보다는 엔트로피를 줄이는 학습이 이루어진다. 또한, MSE의 감소 학습이 둔해지기 시작하면서부터 엔트로피의 감소 학습은 빠르게 진행된다. 따라서, 학습 속도가 부진하여 MSE를 줄이는 작업이 늦어지는 방법은 구조를 최적화시키는 속도도 느려지게 된다.

예로써, SGA와 RSGA는 학습 초기부터 MSE가 GA와 RGA에 비해 빠르게 감소된다. 따라서, GA와 RGA의 경우는 엔트로피 학습이 500~600회 정도부터 이루어지는 것에 비해 SGA와 RSGA의 경우는 200~300회 정도부터 엔트로피 학습이 이루어지며 약 700회 정도에 거의 학습이 종료되어 학습이 빠르게 진행된다.



(a) Reduction of MSE



(b) Reduction of entropy

그림 6. IRIS에 대한 MSE와 엔트로피 감소 과정
Fig. 6. The reduction process of MSE and entropy for IRIS.

4. 각 방법에서의 오인식율 고찰

각 방법에 대한 오인식율 또한 위의 경우들과 마찬가지로이다. IRIS의 경우 각 방법의 학습 패턴에 대한 MSE는 차이가 있으나 오인식율은 거의 유사하다. 그러나, E13B와 LETTER의 경우에는 오인식율에서 차이를 보인다. 이는 표준 편차를 보면 알 수 있듯이 2,000회 반복의 학습 횟수동안 GA와 RGA는 완전하게 학습하지 못하는 경우가 발생하게 되어 평균적으로 오인식율이 높다. 표 2에서 학습이 빠른 시간에 잘 이루어지는 방법일수록 표 5에서 오인식율 또한 낮게 나타났다.

표 5. 오인식율 비교(%)

Table 5. Comparisons of the misclassification ratio(%).

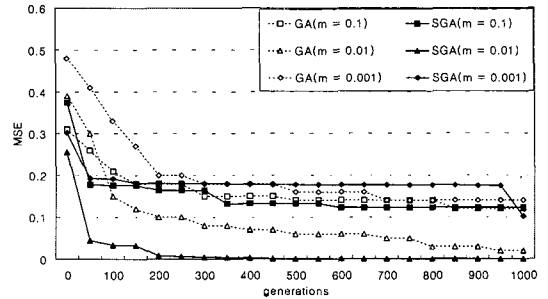
데이터 방법	IRIS	E13B	LETTER
GA	4.5 ± 0.1	6.82 ± 1.45	27.35 ± 5.34
RGA	4.5 ± 0.1	6.77 ± 1.29	26.69 ± 5.16
SGA	4.3 ± 0.1	4.84 ± 0.34	23.15 ± 3.12
RSGA	4.3 ± 0.1	4.79 ± 0.31	22.79 ± 3.14

5. 유전 연산 파라메타 설정에 대한 고찰

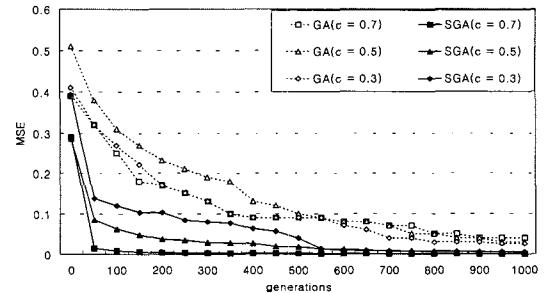
교배, 돌연변이 확률 등 파라메타의 설정은 GA와 SGA의 성능에 영향을 미친다. 우선 돌연변이 확률의 영향을 평가하기 위하여 교배 확률(m)을 0.7로 고정한 후 돌연변이 확률(c)을 0.1, 0.01, 0.001로 실험한 결과는 그림 7의 (a)와 같다. 0.01로 하였을 경우에는 정상적으로 학습을 수행하였으나, 0.1이나 0.001로 하였을 경우에는 학습이 제대로 이루어지지 않았다. 돌

연변이 확률의 변화에 따른 학습 결과는 SGA가 GA에 비해 민감하였다.

교배 확률의 영향을 평가하기 위해서 돌연변이 확률을 0.01로 고정한 후 교배 확률을 0.3, 0.5, 0.7로 하여 실험한 결과는 그림 7의 (b)와 같다. 교배 확률의 경우는 돌연변이 확률과는 달리 그 확률값의 변화가 학습의 성능에 크게 영향을 미치지 못하였다.



(a) Comparisons of MSE on mutation ratio(c = 0.7)



(b) Comparisons of MSE on crossover ratio(m = 0.01)

그림 7. 교배와 돌연변이 확률에 따른 MSE의 비교(교배 확률 = c, 돌연변이 확률 = m)

Fig. 7. Comparisons of MSE on mutation and crossover ratio(crossover ratio = c, mutation ratio = m)

IV. 결론

이 논문에서는 종족 개념을 이용한 GA로 주어진 응용에 적합한 선택적 연결 구조의 MLP 분류기를 효율적으로 생성할 수 있는 학습 방법(SGA)을 제안하였다. 종족을 이용하여 전체 영역을 분할·제거하여 불필요한 영역의 탐색을 줄였으며, 또한 대칭성을 이용한 탐색 공간 축소 방법을 이용하여 보다 적은 학습 시간 내에 불필요한 뉴런 및 링크를 제거한 효율적인 MLP 분류기를 도출하였다.

IRIS, E13B, LETTER의 세 가지 벤치 마크 데이

터에 대하여 학습한 후 테스트한 결과, 종족을 이용한 학습 방법은 기존의 방법에 비해 오인식율, 학습 속도, MLP 분류기의 구조 등에 있어서 효율적임을 확인하였다. 앞으로는 주성분 분석(principal component analysis ; PCA)을 이용하여 입력 뉴런의 수를 최소화한 다음 종족 유전 알고리즘을 적용하는 방법을 연구할 계획이다.

참 고 문 헌

[1] M. Hayashi, "A fast algorithm for the hidden units in a multilayer perceptron," Proceedings of Int. Joint Conf. on Neural Networks, Nagoya, vol. 1, pp. 339-342, 1993.

[2] 押野 隆弘, 尾島 潤, 山本 眞司, "誤差逆伝搬學習における中間層ユニットの 逐次削減法", 電子情報通信學會 論文誌 D-II, vol. J76-D-II, no. 7, pp. 1414-1424, 1993.

[3] K. Murase and Y. Matsunaga, Yoshiaki Nakade, "A back-Propagation algorithm which automatically determines the number of association units", Proceedings of Int. Joint Conf. on Neural Networks, pp. 783-788, 1991.

[4] D. Whitley and C. Bogart, "The Evolution of Connectivity : Pruning Neural Networks Using Genetic Algorithm", Proceedings of Int. Joint Conf. on Neural Networks, vol. 1, pp. 134-137, 1991.

[5] 신성효, 김상운, "GA를 이용한 MLP 분류기의 효율적인 구조 학습", 대한전자공학회 하계종합 학술대회 논문집, vol. 19, no. 1, pp. 258-261, 1996

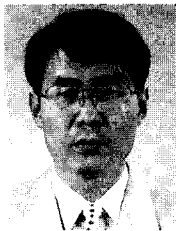
[6] J. C. Potts, T. D. Gidden, and S. B. Yadav, "The Development and Evaluation of an Improved Genetic Algorithm Based on Migration and Artificial Selection", IEEE Trans. on SMC, vol. 24, no. 1, pp. 73-86, Jan. 1994.

[7] 신성효, 김상운, "종족 유전 알고리즘을 이용한 MLP분류기의 학습," 대한전자공학회 하계종합 학술대회 논문집, vol. 20, no. 1, pp. 303-306, 1997

[8] M. Srinivas and L. M. Patnaik, "Learning Neural Network Weights Using Genetic Algorithms-Improving Performance by Search-Space Reduction", Proceedings of Int. Joint Conf. on Neural Networks, pp. 2331-2336, 1991.

[9] M. Ishkawa and H. Uchida, "A Structural Learning of Neural Networks Based on an Entropy Criterion", Proceedings of Int. Joint Conf. on Neural Networks, II 375-380, 1992.

저 자 소 개



金 商 雲(正會員)

1956년 3월 13일생. 1978년 2월 한국항공대학교 통신정보공학과(공학사). 1980년 2월 연세대학교 대학원 전자공학과(공학석사). 1988년 2월 연세대학교 대학원 전자공학과(공학박사). 1984년 4월 ~ 1989년 3월 한

국방송대학교 전자계산학과 조교수. 1992년 12월 ~ 1993년 12월 홋카이도대학 정보공학과 Post-Doc. 1989년 4월 ~ 현재 명지대학교 컴퓨터공학과 부교수. 주관심분야는 패턴인식 및 학습 미디어공학



慎 晟 孝(正會員)

1970년 1월 25일생. 1992년 2월 명지대학교 전자계산학과(공학사). 1994년 2월 명지대학교 대학원 컴퓨터공학과(공학석사). 1994년 3월 명지대학교 대학원 박사과정 입학. 주관심분야는 패턴인식 및 학습, 미디어공

학