다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크 설계 및 최적화

논 문
58-4-25

Design of Fuzzy Relation-based Fuzzy Neural Networks with Multi-Output and Its Optimization

박 건 준* · 김 현 기** · 오 성 권†
(Keon-Jun Park · Hyun-Ki Kim · Sung-Kwun Oh)

Abstract - In this paper, we introduce an design of fuzzy relation-based fuzzy neural networks with multi-output. Fuzzy relation-based fuzzy neural networks comprise the network structure generated by dividing the entire input space. The premise part of the fuzzy rules of the network reflects the relation of the division space for the entire input space and the consequent part of the fuzzy rules expresses three types of polynomial functions such as constant, linear, and modified quadratic. For the multi-output structure the neurons in the output layer were connected with connection weights. The learning of fuzzy neural networks is realized by adjusting connections of the neurons both in the consequent part of the fuzzy rules and in the output layer, and it follows a back-propagation algorithm. In addition, in order to optimize the network, the parameters of the network such as apexes of membership functions, learning rate and momentum coefficient are automatically optimized by using real-coded genetic algorithm. Two examples are included to evaluate the performance of the proposed network.

Key Words : Fuzzy neural networks, Fuzzy relation, Multi-output, Real-coded genetic algorithm, Optimization

1. 서 론

퍼지뉴럴네트워크는 복잡성과 불확실성를 나타내는 비선형 공정 및 패턴분류에서 다른 지능형 모델에 비해 보다 나은 성능을 가지고 있다. 퍼지논리는 1965년 Zadeh[1]에 의해 창안된 "퍼지 집합" 이론을 이용하여 비선형적이고 복잡한 실 시스템의 특성을 해석하는데 적용함으로써 수학적 모델보다 좋은 결과를 가져왔다. 이것은 기존의 0이나 1중 하나의 값만을 선택해야 하는 이분법과는 달리 0과 1사이의 폐구간내 임의의 실수값을 가질 수 있어 비선형 시스템에 내재되어 있는 불확실성을 효과적으로 나타내었으며, 이들 시스템의 정적 혹은 동적 특성을 언어 변수를 사용하여 묘사하였다. 이 언어적인 변수를 표현하는 입력변수, 즉 멤버쉽 함수는 주어진 데이터에 대하여 퍼지 집합으로 표현하여 소속정도를 나타낸다. 하지만 멤버쉽 함수와 규칙을 결정하는데 많은 어려움을 가지고 있다. 이러한 문제를 해결하기 위해 학습능력을 가진 신경망과의 결합이 시도되었다. 하지만 여전히 멤버쉽 함수의 개수, 멤버쉽 함수 중심 등과 신경회로망에서는 층의 개수, 각 층의 노드 개수, 학습율, 모멘텀 계수 등과 같은 초기값 결정에 체계적인 방법이 제시되어 있지 않아 이를 조정하기 위해서 시행착오 및 전문가의 경험으로 결정해야하는 단점이 있다. 또한 이들의 잘못된 선택은 모델자체의 수렴성이나 정확성에 악영향을 미칠 수 있기 때문에 신중한 선택이 요구된다. 따라서, 이들의 선택시 좀 더 객관적인 방법을 이용한다면 모델의 신뢰성을 높일 수 있다[2-5].

퍼지 모델은 시스템을 이해하기 쉽도록 정성적인 표현이 용이하고 불확실한 정보를 가지고 있는 시스템에 강인성을 발휘할 수 있다. 그러나 학습기능을 가지고 있지 않고 있으며 퍼지규칙을 얻기 위해 전문가나 경험에 의존해야 하는 어려움이 있다. 반면에 신경회로망은 시스템의 입출력 관계를 학습을 통해 추출할 수 있으며 병렬처리 기능을 가지고 있어 빠른 처리 능력을 가진다. 그러나 신경회로망은 주어진 시스템에 대해 정성적 표현이 어려워 시스템을 이해하기에는 어려움을 가진다. 이러한 두 모델의 장점을 결합하고 단점을 해결하기 위해 퍼지 모델과 신경회로망의 유사성과 상호 보완관계를 이용하여 다양한 결합을 시도하고 있으며 이러한 구조가 퍼지뉴럴네트워크이다[6-11]. 퍼지뉴럴네트워크는 기존의 모델에 비해서 학습속도가 빠르고, 수렴특성이 우수한 장점을 가진다.

본 논문에서는 전체 입력 공간을 상호 관계된 분할 공간으로 구성하고 보수적 관계를 갖는 멤버쉽 함수를 이용하여 퍼지 규칙의 전반부를 형성한다. 퍼지 규칙의 후반부는 간략 추론, 선형 추론, 변형된 2차식 추론 방법을 사용하여 표현하여 후반부 다항식 계수는 오류 역전파 알고리즘을 이용하여 학습한다. 다중 출력을 가지는 네트워크를 설계하기 위하여 출력층의 노드를 출력수 만큼 생성하며 연결가중치를 이용하여 각 퍼지 규칙에 연결한다. 오류 역전파 알고리즘은 퍼지 규칙내의 후반부 다항식 계수와 출력층의 연결가중치를 학습하는데 사용된다. 설계된 네트워크는 실수코딩

* 정 회 원 : 수원대 공대 전기공학과 박사과정
** 정 회 원 : 수원대 공대 전기공학과 교수 · 공박
† 교신저자, 정회원 : 수원대 공대 전기공학과 교수 · 공박
E-mail : ohsk@suwon.ac.kr
접수일자 : 2008년 10월 13일
최종완료 : 2009년 2월 4일

유전자 알고리즘[12, 13]을 이용하여 각 입력 변수에 대한 전반부 멤버쉽함수의 정점과 학습률 및 모멘텀 계수와 같은 파라미터들을 최적으로 동조한다. 또한, 제안된 네트워크는 다중 출력을 가지는 두 가지의 모의 실험 데이터를 이용하여 패턴 분류 및 평가한다.

2. 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크 설계

퍼지뉴럴네트워크는 퍼지이론과 신경회로망의 두 구조가 유사성과 상호 보완적인 특성으로부터 결합이 시도되었다. 이 결합에 의한 퍼지뉴럴네트워크 구조는 퍼지이론의 언어적 논리를 규칙의 형태로 표현할 수 있는 능력과 신경회로망의 학습기능과의 상호간에 높은 융통성을 가지게 된다.

2.1 퍼지뉴럴네트워크의 구조

입력에 따른 퍼지공간분할은 상호 관계된 분할 공간에 의해 퍼지 규칙을 형성하며 후반부 구조는 간략 추론(Type 0), 선형 추론(Type 1), 변형된 2차식 추론(Type 2)에 의해 표현된다. 각 퍼지 규칙의 추론된 값은 다중 출력을 위해 연결가중치로 연결된 구조이다. 그림 1은 제안된 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크의 구조를 보여준다.

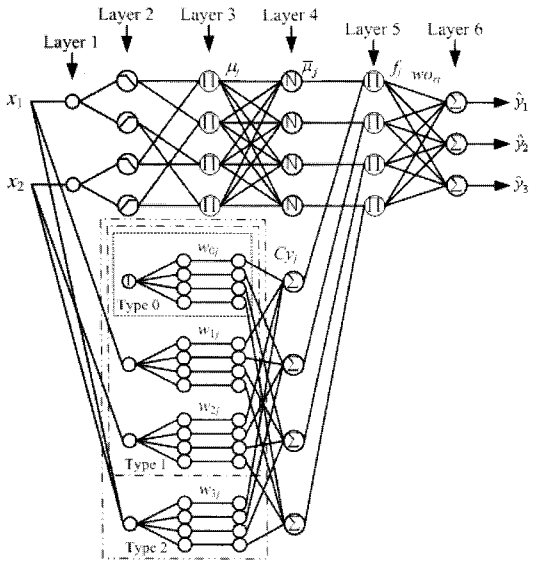


그림 1 퍼지뉴럴네트워크 구조
Fig. 1 Fuzzy Neural Networks structure

그림 1에서 각 입력신호 $x_k(1\le k\le l)$는 각 멤버쉽함수를 거쳐 적합도가 구해지고 적합도는 후반부 출력 Cy_j와 곱해지며 j번째($1\le j\le n$) 규칙의 출력 f_j를 생성한다. r번째($1\le r\le s$) 최종 출력값은 출력층 연결가중치 wo_{rj}와 곱해진 다음 대수적으로 합해지며 다음과 같다.

$$\hat{y}_r=\sum_{j=1}^{n} f_j\cdot wo_{rj} \tag{1}$$

퍼지 규칙은 네트워크의 전체 구조가 퍼지규칙을 이루며, 그 규칙은 다음과 같이 표현된다.

$$R^j:\ If\ x_1\ is\ A_{1j}\text{and}\cdots\text{and}x_k\ is\ A_{kj}\ then\ Cy_j=g(x_1,\cdots,x_k) \tag{2}$$

간략 추론 : $g=w_{0j}$

선형 추론 : $g=w_{0j}+\sum_{k=1}^{l} w_{kj}\cdot x_k$

변형된 2차식 추론 :

$$g=w_{0j}+\sum_{k=1}^{l} w_{kj}\cdot x_k+w^r_{(k+1)j}\cdot x_1\cdot x_2+\cdots +w^r_{(k(k+1)/2)j}\cdot x_{k-1}\cdot x_k$$

여기서, R^j는 j번째 퍼지 규칙을 나타내며, A_{kc}는 전반부 퍼지변수로 멤버쉽함수로 표현된다. w_{kj}는 후반부 계수이며, 퍼지 규칙내의 연결가중치를 나타낸다.

전반부 멤버쉽함수는 이웃하는 멤버쉽함수와 보수적인 관계를 가진 삼각형 멤버쉽함수를 사용한다. 따라서 j번째 규칙에 대한 출력은 다음과 같다.

$$f_j=\overline{\mu_j}\cdot Cy_j \tag{3}$$

제안된 퍼지뉴럴네트워크 구조에서 각 층의 연산과정은 다음과 같다.

[1층] 입력층

각 입력은 퍼지 입력 공간으로부터 분할된 멤버쉽함수로 출력된다.

[2층] 입력에 대한 적합도 계산

멤버쉽함수는 그림 2와 같이 보수적 관계를 갖는 삼각형 멤버쉽함수로, 각 입력에 의해 멤버쉽함수로부터 멤버쉽 값을 구한다.

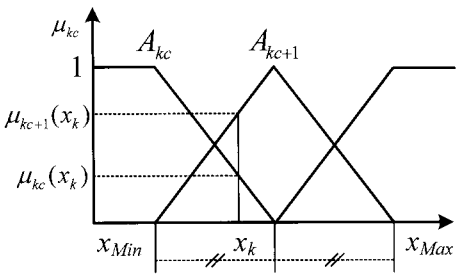


그림 2 보수적 관계를 갖는 멤버쉽함수
Fig. 2 Complementary membership function

[3층] 각 규칙의 적합도 계산

각 노드는 입력되는 신호들을 곱하여 상호 결합된 퍼지 부분공간을 형성하며, 각 규칙의 적합도를 계산한다.

$$\mu_j=\mu_A(x_1)\times\cdots\times\mu_A(x_k) \tag{4}$$

[4층] 정규화된 적합도 계산

각 규칙의 적합도가 구해지면 정규환된 적합도 $\bar{\mu}_{kc}$ 를 계산한다.

$$\bar{\mu}_j = \frac{\mu_j}{\sum_{j=1}^{n} \mu_j} \tag{5}$$

[5층] 각 규칙의 출력

4층에서 구해진 적합도는 후반부 출력과 곱해져 출력층의 입력이 된다.

$$f_j = \frac{\mu_j \cdot Cy_j}{\sum_{j=1}^{n} \mu_j} = \bar{\mu}_j \cdot Cy_j \tag{6}$$

[6층] 최종 출력

최종 출력은 각 규칙의 출력과 출력층의 연결가중치의 가중합에 의해 구해진다.

$$\hat{y}_r = \sum_{j=1}^{n} f_j \cdot wo_{rj} \tag{7}$$

2.2. 퍼지뉴럴네트워크의 학습

제안된 네트워크의 학습은 퍼지규칙내의 연결가중치 w_{kj} 와 출력층의 연결가중치 wo_{rj}를 가장 적절한 값으로 조정해 나가면서 이루어진다. 학습방법은 오류역전파 알고리즘을 이용하며, 이는 원 출력과 모델 출력간의 차, 즉 오차를 바탕으로 이루어진다.

$$E_p = \frac{1}{2} \sum_{r=1}^{s} (y_{pr} - \hat{y_{pr}})^2 \tag{8}$$

여기서, E_p 는 p번째 데이터에 대한 오차이고, y_{pr} 와 $\hat{y_{pr}}$ 는 원 출력과 모델 출력을 나타낸다.

2.1.1 출력층의 연결가중치 조절

출력층의 연결가중치 조절은 다음과 같이 기존의 값에 오차로부터 구해진 조절량 만큼의 변화를 더해 준다.

$$wo_{rj}(t+1) = wo_{rj}(t) + \Delta wo_{rj} \tag{9}$$

오차 E_p 를 최소화하기 위한 연결가중치의 변화량은 경사하강법을 이용하여 구할 수 있다.

$$\Delta wo_{rj} = \eta \left(- \frac{\partial E_p}{\partial wo_{rj}} \right) \tag{10}$$

여기서, η 는 학습률이고, 윗 식은 사슬법칙에 의해 다음과 같이 전개된다.

$$- \frac{\partial E_p}{\partial wo_{rj}} = - \frac{\partial E_p}{\partial \hat{y_{pr}}} \cdot \frac{\partial \hat{y_{pr}}}{\partial wo_{rj}} \tag{11}$$

여기서,

$$- \frac{\partial E_p}{\partial \hat{y_{pr}}} = - \frac{\partial}{\partial \hat{y_{pr}}} \left(\frac{1}{2} \sum_{r=1}^{s} (y_{pr} - \hat{y_{pr}})^2 \right) = y_{pr} - \hat{y_{pr}} \tag{12}$$

$$\frac{\partial \hat{y_{pr}}}{\partial wo_{rj}} = \frac{\partial}{\partial wo_{rj}} \left(\sum_{j=1}^{n} f_j \cdot wo_{rj} \right) = f_j \tag{13}$$

따라서,

$$\Delta wo_{rj} = \eta (y_{pr} - \hat{y_{pr}}) f_j \tag{14}$$

또한, 모멘텀을 추가해 학습시간을 좀 더 단축시키고, 학습성능을 향상시킨다. 연결가중치의 변화량이 클 경우에는 변화량을 키워주고, 변화량이 작아지면 변화량을 줄여 주는 역할을 한다. 모멘텀은 현재의 연결가중치와 그 이전 연결가중치의 차이에 의해 계산된다. 이렇게 계산된 모멘텀은 모멘텀 계수 α와 곱해져 연결 가중치 조절식에 추가된다. 모멘텀까지 포함하는 연결 가중치의 최종 조절식은 다음과 같다.

$$\Delta wo_{rj} = \eta (y_{pr} - \hat{y_{pr}}) f_j + \alpha \mathrm{wo_{rj}(t) - wo_{rj}(t-1)} \tag{15}$$

2.1.1 퍼지규칙내의 연결가중치 조절

퍼지규칙내의 후반구 다항식 계수인 연결가중치의 조절은 출력층의 연결가중치에서와 같은 방법으로 이루어진다.

간략 추론 방법인 경우는 다음과 같다.

$$w_{0j}(t+1) = w_{0j}(t) + \Delta w_{0j} \tag{16}$$

$$\Delta w_{0j} = \eta \left(- \frac{\partial E_p}{\partial w_{0j}} \right) \tag{17}$$

$$- \frac{\partial E_p}{\partial w_{0j}} = \sum_{r=1}^{s} \left(- \frac{\partial E_p}{\partial \hat{y_{pr}}} \cdot \frac{\partial \hat{y_{pr}}}{\partial f_j} \right) \cdot \frac{\partial f_j}{\partial w_{0j}} \tag{18}$$

여기서,

$$- \frac{\partial E_p}{\partial \hat{y_{pr}}} = - \frac{\partial}{\partial \hat{y_{pr}}} \left(\frac{1}{2} \sum_{r=1}^{s} (y_{pr} - \hat{y_{pr}})^2 \right) = y_{pr} - \hat{y_{pr}} \tag{19}$$

$$\frac{\partial \hat{y_{pr}}}{\partial f_j} = \frac{\partial}{\partial f_j} \left(\sum_{j=1}^{n} f_j \cdot wo_{0j} \right) = wo_{0j} \tag{20}$$

$$\frac{\partial f_j}{\partial w_{0j}} = \frac{\partial}{\partial w_{0j}} (\mu_j \cdot w_{0j}) = \mu_j \tag{21}$$

따라서,

$$\Delta w_{0j} = \eta \sum_{r=1}^{s} \{ (y_{pr} - \hat{y_{pr}}) wo_{0j} \} \mu_j \tag{22}$$

또한, 모멘텀 계수를 포함하는 연결가중치 변화량은 다음과 같다.

$$\Delta w_{0j} = \eta \sum_{r=1}^{s} \{(y_{pr} - \hat{y_{pr}}) wo_{0j}\} \mu_j + \alpha w_{0j}(t) - w_{0j}(t-1) \quad (23)$$

선형 추론인 경우와 변형된 2차식 추론인 경우에도 같은 방법으로 유도되며, 각각 다음과 같다.

$$\Delta w_{kj} = \eta \sum_{r=1}^{s} \{(y_{pr} - \hat{y_{pr}}) wo_{0j}\} \mu_j x_k + \alpha w_{kj}(t) - w_{kj}(t-1) \quad (24)$$

$$\Delta w_{(k(k+1)/2)j} = \eta \sum_{r=1}^{s} \{(y_{pr} - \hat{y_{pr}}) wo_{0j}\} \mu_j x_{k-1} x_k + \alpha w_{(k(k+1)/2)j}(t) - w_{(k(k+1)/2)j}(t-1) \quad (25)$$

3. 실수코딩 유전자 알고리즘을 이용한 최적화

유전자 알고리즘[12, 13]이란 적자생존의 생물학 원리에 바탕을 둔 최적화 기법중의 하나로 자연계의 생명체 중 환경에 잘 적응한 개체가 좀 더 많은 자손을 남길 수 있다는 자연선택 과정과 자연계 생명체의 설계도와 같은 유전자의 변화를 통해서 좋은 방향으로 발전해 나간다는 자연 진화 과정의 유전자 메커니즘에 바탕을 둔 탐색 알고리즘이다. 즉, 자연계에 있어서 생물학의 유전과 진화의 메커니즘을 공학적으로 모델화하는 것에 의해 생물의 환경 적응능력과 자연도태의 원리를 기반으로 만들어진 계산 모델로, 실세계의 문제를 풀기 위해 잠재적인 해들을 컴퓨터상에서 코딩된 개체로 나타내고, 여러 개의 개체들은 모아 군집을 형성한 뒤, 세대를 거듭하면서 이들의 유전 정보를 서로 교환하거나 새로운 유전 정보를 부여하면서 적자생존의 법칙에 따라 모의 진화를 시킴으로써, 주어진 문제에 대한 최적의 해를 찾는 계산 모델이다.

유전자 알고리즘은 기존의 다른 탐색방법들과 달리 유전자 알고리즘은 해가 될 가능성이 있는 개체집단을 유지하면서 그들 모두가 동시에 최적 값을 찾아나가기 때문에 지역 극소에 빠질 위험을 해결할 수 있다는 점과 더불어 모델의 성능지수가 최소가 되는 전역 극소 영역을 찾는 능력을 갖고 있으며, 또한 기존의 방법들과는 달리 선형, 연속, 미분가능 등의 제한이 없기 때문에 다양한 분야에 적용할 수 있다는 장점을 가진다는 것이 중요한 특징이다.

유전자 알고리즘은 3개의 기본 연산자, 즉 재생산, 교배, 돌연변이를 가지고 있다. 생산(reproduction)은 다음 세대를 위해 교배 유전자가 선택되어지는 과정으로, 적합도에 따라 다음세대를 생산할 수 있는 기회를 부여한다. 각각의 문자열들은 그들의 적합도를 가지고 다음세대에 전달된다. 교배(crossover)는 교배 유전자로부터 염색체들은 짝짓고 짝지어진 문자열 쌍은 확률적으로 선택되어진다. 돌연변이(mutation)는 문자열 위치의 값을 임의로 변경한다. 2진 코딩에서 돌연변이는 0을 1로 또는 1을 0으로 바꾸는 것을 의미한다. 적절하게 정의된 적합도 함수와 함께 조합된 연산자는 유전자 알고리즘의 주요 부분이 된다. 유전자 알고리즘은 초기 집단 생성후 적합도를 평가하고, 재생산, 교배, 돌연변이의 기본 연산자를 통해 최적해를 탐색하며, 일반적인 수행과정은 그림 3과 같다.

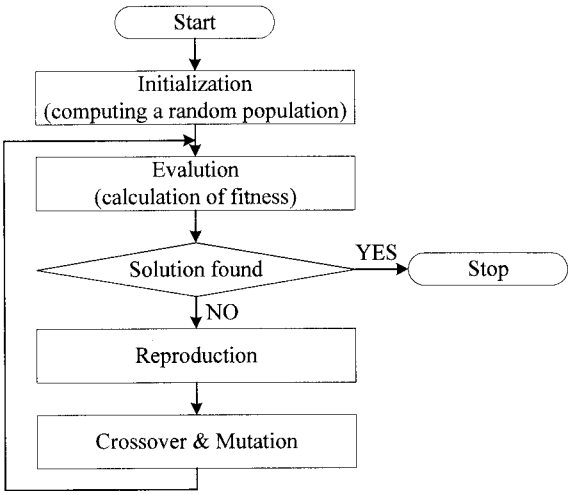


그림 3 유전자 알고리즘의 수행과정
Fig. 3 The procedure of the genetic algorithms

실세계의 문제들은 고정밀도의 해를 요구하고, 해에 대한 사전지식이 결여되거나, 복잡한 제약조건이 존해하는 경우가 종종 있다. 이진코딩을 채용할 때 정밀도를 높이거나 해에 대한 사전지식의 결여로 탐색구간을 확대하게 되면 염색체 길이가 길어진다. 긴 염색체는 매우 큰 탐색공간을 만들어 해를 찾는데 많은 시간을 요구하거나, 경우에 따라서는 실패하는 원인을 제공한다. 이러한 문제점들은 근본적으로 염색체 표현을 좀 더 해공간에 접근시켜 문제와 관련된 지식을 포함시키는 연산자 설계를 용이하게 하고, 제약조건을 다루기 위한 도구를 설계하는 것이 쉽도록 함으로써 해결할 수 있으며, 대표적인 표현법이 실수코딩이다. 실수코딩을 채용함으로써 염색체의 유전자와 벡터 요소는 정확히 일대일 일치하는 관계를 가지며 염색체의 길이는 벡터의 차원과 같게 된다.

본 논문에서는 제안된 네트워크에 대한 전반부 멤버쉽함수의 정점과 학습률 및 모멘텀 계수를 최적으로 설정해 주기 위하여 실수 코딩 유전자 알고리즘을 이용하여 네트워크의 최적화 과정을 수행하게 된다. 그림 4는 네트워크의 최적화를 위해 첫 번째 입력 변수 x_1에서 k번째 입력 변수 x_k까지 3개의 멤버쉽함수를 할당하고 그 다음으로 학습률 및 모텀텀 계수를 일련의 문자열로 할당한 염색체 구조를 보여준다. 이들 염색체들은 유전 연산자에 의해 세대를 거듭하면서 최적 해를 탐색한다.

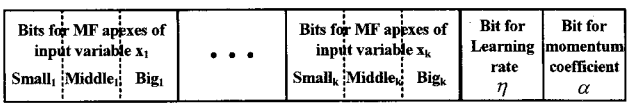


그림 4 염색체 구조
Fig. 4 The structure of chromosomes

4. 실험 데이터를 통한 결과 고찰

본 논문에서 제안된 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크를 적용하기 위하여 다중 패턴 실험 데이터로 Synthetic Data와 Fisher의 Iris Data를 이용하였다.

4.1 Synthetic Data

본 논문에서 제안된 퍼지뉴럴네트워크를 적용하기 위하여 2-클래스를 가진 모의 데이터 집합을 구성하여 실험하였다. 데이터 집합은 각 클래스별로 150개의 패턴을 가진, 300개의 패턴으로 구성된다. 그림 5는 구성된 데이터 집합의 분포를 보여준다.

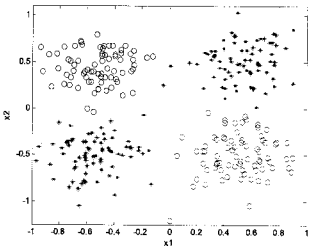


그림 5 2-클래스 모의 데이터 집합
Fig. 5 2-class synthetic data set

제안된 네트워크를 이용하여 패턴분류를 수행하였으며, 실험은 random sub-sampling을 이용하여 각 데이터 집합을 7:3의 비율로 랜덤하게 학습 데이터와 테스트 데이터로 나누어 5번 반복 실험하였다. 성능평가는 오차 및 MSE(Mean Square Error)에 의한 정확도의 평균과 표준편차에 의해 평가하였다.

$$MSE=\frac{1}{m}\frac{1}{s}\sum_{p=1}^{m}\sum_{r=1}^{s}(y_{pr}-\hat{y_{pr}})^2 \qquad (26)$$

퍼지뉴럴네트워크에 의한 패턴분류는 각 입력 데이터의 최소값과 최대값을 균등 분할하여 전체 입력 공간 분할을 시행하였고, 공간 분할에 따른 전반부 멤버쉽함수의 정점을 각 입력 변수에 대해 2개 또는 3개씩 할당하였으며, 후반부 구조로는 간략 추론, 선형 추론, 변형된 2차식 추론을 사용하였다. 후반부 다항식 계수 및 출력층의 연결가중치는 오류 역전파 알고리즘을 이용하여 학습하였다. 또한, 실수코딩 유전자 알고리즘을 적용하여 전반부 멤버쉽함수의 정점과 학습률, 모멘텀 계수를 동조하였다.

표 1은 퍼지뉴럴네트워크 및 유전자 알고리즘의 파라미터를 보여준다. 퍼지뉴럴네트워크에서 학습은 300번 반복하였으며, 멤버쉽함수의 수는 각 입력 변수에 대하여 2개인 경우와 3개인 경우를 나누어 각각 실행하였다. 학습률 및 모멘텀 계수는 각각 0.01과 0.001로 설정하였다. 실수코딩 유전자 알고리즘 기반 퍼지뉴럴네트워크에서는 세대수 150, 개체수 60로 설정하였다. 재생산 연산자는 룰렛휠 선택을 사용하였으며, 교배연산자는 수정단순교배를 사용하고 교배율은 0.65로 하였다. 돌연변이 연산자와 돌연변이율은 각각 단순돌연변이와 0.1로 하였다. 또한, 집단 내에서 최적 개체를 보존하기 위하여 엘리트 전략을 사용하였다.

표 1 퍼지뉴럴네트워크 및 유전자 알고리즘의 파라미터
Table 1 Parameters for FNN and GAs

(a) 퍼지뉴럴네트워크

	파라미터	값
Fuzzy Neural Network (FNN)	학습횟수	300
	멤버쉽함수 (MF)의 수	각 입력당 2개 / 3개
	학습률	0.01
	모멘텀 계수	0.001

(b) 퍼지뉴럴네트워크 및 유전자 알고리즘

	파라미터	값
Fuzzy Neural Network (FNN)	학습횟수	150
	멤버쉽함수 (MF)의 수	각 입력당 2개 / 3개
Genetic Algorithm (GA)	세대수	150
	개체수	60
	교배율	0.65
	돌연변이율	0.1

표 2는 각 입력 변수에 대해 멤버쉽함수의 수가 2개인 경우와 3개인 경우에 있어서 후반부 구조가 간략 추론, 선형 추론, 변형된 2차식 추론방법에 의한 퍼지뉴럴네트워크 및 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 성능을 보여준다.

표 2 성능평가
Table 2 Performance evaluation

(1) 퍼지뉴럴네트워크의 성능 평가

FNN		인식오차		정확도	
MF 수	추론	학습	테스트	학습	테스트
2	간략	3.60±1.52	3.00±2.00	0.063±0.00	0.097±0.03
	선형	3.80±2.17	3.80±2.59	0.065±0.00	0.088±0.01
	2차식	4.80±2.39	2.80±2.17	0.064±0.00	0.076±0.01
3	간략	2.20±1.10	8.60±3.51	0.058±0.01	0.098±0.02
	선형	3.00±1.00	3.20±1.92	0.046±0.00	0.062±0.01
	2차식	3.00±1.00	6.00±3.94	0.039±0.01	0.070±0.02

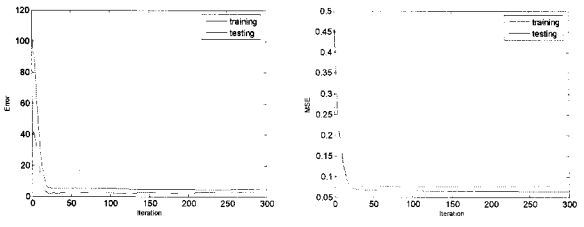
(2) 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 성능 평가

GA+FNN		인식오차		정확도	
MF 수	추론	학습	테스트	학습	테스트
2	간략	0.60±1.34	1.20±1.30	0.032±0.01	0.037±0.02
	선형	0.40±0.55	0.20±0.45	0.018±0.01	0.022±0.01
	2차식	0.80±1.10	1.40±1.52	0.015±0.01	0.031±0.01
3	간략	0.40±0.55	0.80±0.84	0.042±0.00	0.051±0.01
	선형	0.40±0.89	2.00±2.92	0.024±0.01	0.040±0.03
	2차식	0.80±1.30	1.00±1.41	0.028±0.01	0.043±0.02

표 2에서 보는 바와 같이 각 입력 데이터의 최소값과 최대값을 이용하여 공간분할을 하고 학습을 한 경우 보다 그

정점을 초기치로 두고 그 정점과, 학습률 및 모멘텀 계수를 유전자 알고리즘으로 동조한 경우 오차와 정확도에 있어서 근사화 및 일반화 모두 성능 향상을 꾀할 수 있었다.

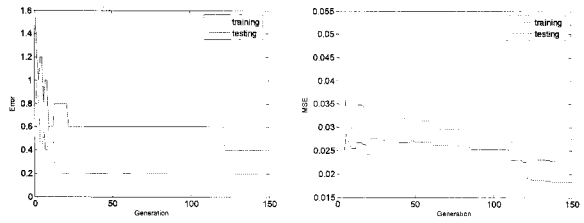
그림 6과 그림 7는 5번 반복에 따른 오차와 정확도의 평균값을 보여준다. 그림 6은 멤버쉽함수의 수가 2개 이고 후반부 구조가 변형된 2차식 추론일 때의 학습에 따른 퍼지뉴럴네트워크의 학습과정을 보여주며, 그림 7는 멤버쉽함수의 수가 2개이고 후반부 구조가 선형추론인 경우, 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 최적화 과정을 보여준다.



(a) 오차 (b) MSE

그림 6 퍼지뉴럴네트워크의 학습 과정

Fig. 6 Learning process of FNN



(a) 오차 (b) MSE

그림 7 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 최적화 과정

Fig. 7 Optimization process of FNN by GAs

4.2. Fisher의 Iris 데이터

제안된 퍼지뉴럴네트워크를 적용하기 위하여 다중 패턴 두 번째 실험 데이터로 Fisher의 Iris 데이터를 이용하였다. 실험에 사용된 Iris 데이터는 3개의 클래스(setosa, versicolor, virginica)로 구성되며, 각 클래스마다 50개의 패턴로 구성되어 있고, 각 패턴은 4개의 속성으로 정의된다.

Fisher의 Iris 데이터를 이용한 퍼지뉴럴네트워크에 의한 패턴분류는 앞선 모의 실험 데이터를 이용한 방법과 같은 방법으로 수행하였으며, 표 1에서 퍼지뉴럴네트워크 및 유전자 알고리즘의 파라미터를 보여준다.

표 3은 각 입력 변수에 대해 멤버쉽함수의 수가 2개인 경우와 3개인 경우에 있어서 후반부 구조가 간략 추론, 선형 추론, 변형된 2차식 추론방법에 의한 퍼지뉴럴네트워크 및 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 성능을 보여준다.

표 3에서 보여주는 것과 같이 각 입력 데이터의 최소값과 최대값을 균등분할하여 학습을 한 경우 보다 유전자 알고리즘을 적용하여 전반부 멤버쉽함수의 정점과 학습률, 모멘텀 계수를 동조한 경우 오차와 정확도에 있어서 근사화 및 일반화 모두 성능이 향상된 것을 알 수 있다.

그림 8과 그림 9는 5번 반복에 따른 오차와 정확도의 평균값을 보여준다. 그림 8은 멤버쉽함수의 수가 2개 이고 후반부 구조가 선형추론일 때의 학습에 따른 퍼지뉴럴네트워크의 학습과정을 보여주며, 그림 9는 멤버쉽함수의 수가 3개 이고 후반부 구조가 변형된 2차식 추론인 경우, 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 최적화 과정을 보여준다.

표 3 성능평가

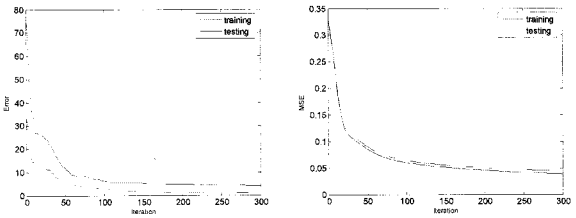
Table 3 Performance evaluation

(1) 퍼지뉴럴네트워크의 성능 평가

FNN		인식오차		정확도	
MF 수	추론	학습	테스트	학습	테스트
2	간략	4.60±1.14	2.60±1.67	0.045±0.00	0.055±0.01
	선형	4.20±0.84	1.00±0.71	0.038±0.00	0.044±0.01
	2차식	2.60±0.89	2.80±1.10	0.034±0.00	0.050±0.01
3	간략	3.60±1.82	1.80±1.48	0.033±0.00	0.041±0.01
	선형	3.80±1.48	2.40±2.79	0.028±0.00	0.043±0.03
	2차식	3.20±1.30	1.00±1.22	0.025±0.00	0.038±0.01

(2) 퍼지뉴럴네트워크 및 유전자 알고리즘에 의한 성능 평가

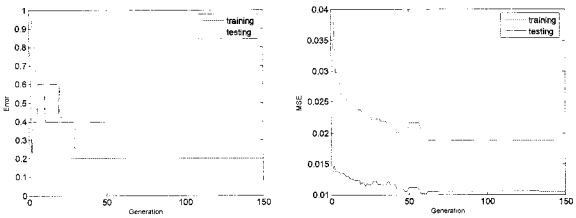
GA+FNN		인식오차		정확도	
MF 수	추론	학습	테스트	학습	테스트
2	간략	0.20±0.45	1.20±0.84	0.019±0.00	0.026±0.01
	선형	0.60±0.55	0.60±0.55	0.021±0.00	0.024±0.01
	2차식	0.60±0.55	1.20±0.45	0.020±0.00	0.026±0.01
3	간략	0.20±0.45	0.40±0.55	0.016±0.00	0.026±0.02
	선형	0.00±0.00	0.40±0.89	0.013±0.00	0.017±0.01
	2차식	0.20±0.45	0.20±0.45	0.010±0.00	0.018±0.01



(a) 오차 (b) MSE

그림 8 퍼지뉴럴네트워크의 학습 과정

Fig. 8 Learning process of FNN



(a) 오차 (b) MSE

그림 9 유전자 알고리즘에 의한 퍼지뉴럴네트워크의 최적화 과정

Fig. 9 Optimization process of FNN by GAs

표 4는 기존 모델과의 비교를 보여준다. 제안된 네트워크는 멤버쉽함수의 수가 3개이고 후반분 구조가 변형된 2차식 추론인 경우에 있어서 학습 데이터와 테스트 데이터에 대한 평균 분류율을 나타내었다. 기존 모델과 비교해 보면 제안된 네트워크는 학습에 의한 근사화 및 테스트에 의한 일반화 성능에 있어서 모두 좋은 성능을 보여주는 것을 알 수 있다.

표 4 기존 모형과의 비교

모델	평균 분류율(%)		
	전체	학습	테스트
Simpson[14]	97.3		
Nauck[15]		96	97.3
		97.3	96
Meesad[16]	88~97.3		
Yen[17]		96.7	96.2
Li[18]	96.67		
Fuzzy Neural Networks		99.8	99.8

5. 결 론

본 논문에서는 다중 출력을 가지는 퍼지 관계 기반 퍼지 뉴럴네트워크를 설계하였으며, 그 설계과정을 상세히 설명하였다. 퍼지 규칙의 전반부는 전체 입력공간을 상호 관계에 의해 분할하여 형성하였으며, 퍼지 규칙의 후반부는 다항식을 이용하여 표현하고 오류 역전파 알고리즘을 이용하여 각 규칙의 후반부 다항식 계수를 학습하였다. 다중 출력을 갖는 네트워크를 설계하기 위하여 출력층의 노드을 출력수 만큼 생성하며 연결가중치를 이용하여 각 퍼지 규칙에 연결하였고 오류 역전파 알고리즘을 이용하여 각 연결가중치를 학습하였다. 또한, 연산시간을 단축하고 염색체 표현을 실수로 표현함으로써 해공간에 좀더 접근시키고 그에 상응하는 유전 연산자를 이용하는 실수코딩 유전자 알고리즘을 이용하여 전반부 멤버쉽함수의 정점과 학습률, 모멘텀 계수를 최적화하였다. 제안된 다중 출력 네트워크는 Synthetic Data와 Fisher의 Iris Data를 이용하여 패턴분류에 적용하였다.

두 가지의 실험 데이터를 통해 모델을 평가하는 데 있어서 오차와 정확도는 서로 밀접한 관계가 있는 것을 알 수 있었다. 즉, 오차가 좋으면 정확도도 좋은 경향을 보여준다. 하지만 오차가 가장 적은 모델이라 할지라도 정확도는 그렇하지 않는 경우가 있으며 반대의 경우도 또한 존재한다. 또한, 비슷한 오차를 갖는 경우에서도 정확도는 서로 차이가 있을 것을 알 수 있다. 따라서, 오차와 정확도 간의 상호보완적인 해석이 요구된다.

실험을 통해 각 입력 데이터의 최소값과 최대값을 이용하여 입력 공간을 균등 분할하여 학습을 한 경우와 그 정점과, 학습률 및 모멘텀 계수를 유전자 알고리즘으로 동조한 경우를 비교해 보면 오차와 정확도에 있어서 근사화 및 일반화 모두 상당한 성능 향상을 보여주며 최적 퍼지뉴럴네트워크를 설계할 수 있었다. 본 연구에 의해 개발된 다중출력 퍼지뉴럴네트워크는 모델링 및 인식율해석에 보다 광범위하게 적용될 것이다.

감사의 글

이 논문은 2007년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음 (KRF-2007-511-D00103, KRF-2008-314-D00376)

참 고 문 헌

[1] L. A Zadeh, "Fuzzy sets", Inf. Control 8, 338-353, 1965.

[2] H. Takagi and I. Hayashi, "NN-driven Fuzzy Reasoning", Int. J. of Approximate Reasoning, Vol. 5, No. 3, pp. 191-212, 1991.

[3] N. Imasaki, J. Kiji, and T. Endo, "A Fuzzy Rule Structured Naural Networks", Journal of Japan Society for Fuzzy Theory and Systems, Vol. 4, No. 5, pp. 987-995, 1992(in Japanese).

[4] H. Nomura and Wakami, "A Self-Tuning Method of Fuzzy Control by Descent Methods", 4th IFSA'91, pp. 155-159, 1991.

[5] Shin-ichi Horikawa, Takeshi Furuhashi, and Yoshiki Uchigawa, "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back Propagation Algorithm", IEEE trans. Neural Networks, Vol. 3, No. 5, pp. 801-806, 1992.

[6] Takeshi Yamakawa, "A Neo Fuzzy Neuron and Its Applications to System Identification and Prediction of the System Behavior", Proceedings of the 2nd International Conference on Fuzzy logic & Neural Networks, pp. 477-483, 1992.

[7] Takeshi Yamakawa, "A New Effectie Learning Algorithm for a Neo Fuzzy Neuron Model", 5th IFSA World Congress, pp. 1017-1020, 1993.

[8] J. J. Buckley and Y. Hayashi, "Fuzzy neural networks: A survey", *Fuzzy Sets and Systems*, vol. 66, pp. 1-13, 1994.

[9] M. M. Gupta and D. H. Rao, "On the principles of fuzzy neural net-works", *Fuzzy Sets and Systems*, vol. 61, pp. 1-18, 1994.

[10] J.-S. R. Jang, C. -T. Sun, E. Mizutani, Neuro-Fuzzy And Soft Computing, Prentice-Hall, 1997.

[11] 오성권, 프로그래밍에 의한 컴퓨터지능, 내하출판사, 2003.

[12] D. E. Golderg, *Genetic Algorithm in search, Optimization & Machine Learning*, Addison wesley, 1989.

[13] 진강규, *유전알고리즘과 그 응용*, 교우사, 2004.

[14] P. K. Simpson, "Fuzzy Min-Max Neural Networks -Part 1 : Classification", *IEEE Trans on. Neural Networks*, vol. 3, no. 5, 1992.

[15] D. Nauck, R. Kruse, "A neuro-fuzzy method to learn

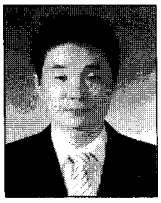
fuzzy classification rules from data", *Fuzzy Sets and Systems*, vol. 89, pp. 277-288, 1997.

[16] P. Meesad, G. G. Yen, "Pattern classification by a neurofuzzy network : application to vibration monitoring", *ISA Transactions*, vol. 39, pp. 293-308, 2000.

[17] G. G. Yen, P. Meesad, "An effective neuro-fuzzy paradigm for machinery condition health monitoring", *IEEE Trans. on Systems, Man, and Cybern.*, : Part B, vol. 31, no. 4, 2001.

[18] R. Li, M. Mukaidono and I.B. Turksen, "A fuzzy neural network for pattern classification and feature selection", Fuzzy Sets and Systems, Vol. 130, pp. 101-140, 2002.

저 자 소 개



박 건 준 (朴 建 俊)

2003년 : 원광대학교 전기전자공학부 졸업. 2005년 : 동 대학원 제어계측공학과 졸업(공학석사). 2005~현재 : 수원대학교 전기공학과 박사과정.

관심분야 : 컴퓨터 및 인공지능, 퍼지추론시스템, 신경망, 유전자 알고리즘 및 최적화이론, 자동화시스템 및 제어 등

Tel : (031) 222-6544

Fax : (031) 220-2667

E-mail : birds@suwon.ac.kr

김 현 기 (金 賢 起)

1977년 : 연세대학교 전기공학과 졸업. 1985년 : 동 대학원 전기공학과 졸업(공학석사). 1991년 : 동 대학원 졸업(공학박사). 1997년~현재 : 경기지역 산학연 기술지원단 단장. 1998년~현재 : 산업기술정책연구소 심의위원. 2000년~현재 : 사단법인 산학연 컨소시엄 전국협의회 회장. 1989년~현재 : 수원대학교 전기공학과 교수. 관심분야 : 시스템자동화 및 감시제어, 지능형모델링 및 제어 등.

Tel : (031) 220-2666

Fax : (031) 220-2667

E-mail : hkkim@suwon.ac.kr

오 성 권 (吳 聖 權)

1981년 : 연세대학교 전기공학과 졸업. 1983년 : 동 대학원 전기공학과 졸업(석사). 1983~1989년 : 금성산전연구소(선임연구원). 1993년 : 연세대 대학원 전기공학과 졸업(공박). 1996~1997년 : 캐나다 Manitoba대학 전기 및 컴퓨터공학과 Post-Doc. 1993~2005년 : 원광대 전기전자 및 정보공학부 교수 2005년~현재 : 수원대 전기공학과 교수. 2002년~현재 : 대한전기학회, 퍼지및지능시스템학회 및 제어자동화시스템 공학회 편집위원.

Tel : (031) 229-8162

Fax : (031) 220-2667

E-mail : ohsk@suwon.ac.kr