

# 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기 설계

## The Design of Pattern Classification based on Fuzzy Combined Polynomial Neural Network

노 석 범\* · 장 경 원\* · 안 태 천†

(Seok-Beom Rho · Kyung-Won Jang · Tae-Chon Ahn)

**Abstract** - In this paper, we propose a fuzzy combined Polynomial Neural Network(PNN) for pattern classification. The fuzzy combined PNN comes from the generic TSK fuzzy model with several linear polynomial as the consequent part and is the expanded version of the fuzzy model. The proposed pattern classifier has the polynomial neural networks as the consequent part, instead of the general linear polynomial. PNNs are implemented by stacking the simple polynomials dynamically. To implement one layer of PNNs, the various types of simple polynomials are used so that PNNs have flexibility and versatility. Although the structural complexity of the implemented PNNs is high, the PNNs become a high order-multi input polynomial finally. To estimate the coefficients of a polynomial neuron, The weighted linear discriminant analysis. The output of fuzzy rule system with PNNs as the consequent part is the linear combination of the output of several PNNs. To evaluate the classification ability of the proposed pattern classifier, we make some experiments with several machine learning data sets.

**Key Words** : Polynomial neural networks, Weighted linear discriminant analysis, Polynomial neuron, Fuzzy combined polynomial neural networks, Discriminant function

### 1. 서 론

패턴 분류기는 정보생성과정을 거쳐 엄청난 양의 데이터들이 발생하는 현시대에 의료, 경제, 고장 진단과 같은 여러 유용한 방면에 사용되어지고 있다. 패턴 분류 기법은 다양한 접근 방법으로 시도되고 있으며, 크게 통계적 기법과 뉴럴 네트워크를 이용한 방법, 규칙 기반 분류 기법으로 분류되어질 수 있다 [1]. 위와 같은 접근 방법들 중에서 규칙 기반 분류 기법은 인간의 언어적 지식을 정량적으로 표현 할 수 있는 규칙으로 표현 할 수 있다는 장점을 가지고 있다. 특히 L. A. Zadeh에 의해 제안된 퍼지 논리는 복잡하고 비선형성이 강하게 내포되어 있는 시스템의 분석 분야에서 좋은 성능을 보이고 있다 [4]. 이와 같은 비선형성에 대한 대처 능력을 가진 퍼지 집합을 기반으로 하는 퍼지 규칙 기반 분류 기법은 인간이 가진 지식의 애매모호함을 표현 할 수 있다는 장점으로 다양한 응용분야에서 널리 사용되고 있다. 다양한 퍼지 규칙 기반 시스템들 중에서 TSK 퍼지 모델은 퍼지 규칙의 후반부가 선형 식으로 정의된 구조를 가지고 있다. TSK 퍼지 모델은 비선형 모델링 분야와 제어기 분야에 적용되어 우수한 성능을 보이고 있는 퍼지 규칙 기반 시스템이다. 그러나 TSK 퍼지 모델의 후반부로 사용되는 선형 다항식은 해공간의 비선형성이 증가할 경우, 선형 다항식

만으로 퍼지 집합으로 정의된 부해(sub solution) 공간에서 입력력 관계를 모델링하는 모델로 적절한 역할을 수행 할 수 없게 된다. 앞서 언급한 TSK 퍼지 모델의 후반부 다항식이 선형식이라는 제한적인 특성을 극복하기 위하여 다양한 다항식들을 가진 다항식 퍼지 모델이 제안되었다[5]. 또 다른 해결 방법으로는 퍼지 규칙이 정의된 부분 공간의 입력 특성을 최대한 모델링하기 위해서 선형 식으로 표현이 부족할 경우 일반적으로 입력 공간을 더 작은 부분 공간으로 나누어 표현한다. 그러나 이와 같은 방법은 전체 입력 공간을 표현하기 위하여 퍼지 규칙의 수를 증가 시켜야 하는 단점을 가지고 있다.

본 논문에서는 퍼지 규칙의 수를 증가 시키지 않고 정의된 퍼지 규칙의 부분 공간의 특성을 표현하기 위하여 퍼지 규칙 기반 시스템의 후반부 구조를 다양화 한 퍼지 결합 다항식 뉴럴 네트워크를 이용한 패턴 분류기를 제안한다. 제안된 퍼지 결합 다항식 기반 패턴 분류기는 TSK 퍼지 규칙 기반 시스템의 확장된 형태로 제안된 퍼지 규칙 기반 시스템의 후반부 구조의 다양성을 확보하기 위하여 동적 생성 네트워크인 다항식 뉴럴 네트워크 [2, 10]을 퍼지 규칙 기반 시스템의 후반부 구조로 사용한다. 다항식 뉴럴 네트워크는 4가지 형태의 구조를 가진 다항식을 기본 노드로 하여 계층적으로 쌓아 올려 만든 네트워크이다. 다항식 뉴럴 네트워크의 최종 식은 일반적인 복잡한 형태의 다항식과 유사하다.

일반적인 다 입력 고차 다항식을 퍼지 규칙 기반 시스템의 후반부 구조로 사용할 경우 다항식의 파라미터 수의 증가로 인하여 주어진 학습 데이터의 수가 충분하지 못 할 경우 추정된 파라미터의 정확도가 떨어진다. 그러나 다항식 뉴럴 네트워크의 경우 비교적 단순한 형태의 다항식을 반복

† Corresponding Author: Dept. of Electronics Convergence Engineering, Wonkwang University, Korea  
E-mail : tcahn@wku.ac.kr

\* Dept. of Electronics Convergence Engineering, Wonkwang University, Korea

Received : December 17, 2013; Accepted : March 6, 2014

적으로 중첩시킴으로서 비선형성이 강한 구조를 설계 할 수 있다는 장점을 가지고 있다. 다항식 뉴럴 네트워크는 A. G. Ivakhnenko가 제안한 Group Method of Data Handling (GMDH) [6] 알고리즘에 바탕을 두고 확장된 네트워크 이다[2].

공간 분할된 각각의 부 공간에 정의된 다항식 뉴럴 네트워크는 퍼지 규칙에 의해 결합되어진다. 그리하여, 제안된 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기의 최종 출력은 각 퍼지 규칙의 적합도와 해당 다항식 뉴럴 네트워크들의 곱의 합으로 정의된다. 다항식 뉴럴 네트워크의 계수를 결정하기 위하여 패턴 분류 분야에서 많이 사용되는 Linear Discriminant Analysis(LDA) 기법을 확장한 Weighted LDA (WLDA)를 사용하였다.

본 논문에서 제안된 패턴 분류기의 패턴 분류 성능을 비교, 평가하기 위하여 여러 개의 기계 학습 데이터를 이용하여 기존의 패턴 분류기와 성능 비교를 수행한다.

## 2. 퍼지 규칙의 후반부 구조 지역적 학습

일반적인 TSK 퍼지 추론 시스템의 후반부 다항식의 계수는 전체 입력 공간에서 추정된다. 그러나 TSK 퍼지 추론 시스템을 이루는 퍼지 규칙들 사이의 상호 연관성을 배제할 수 있다면, 퍼지 규칙의 후반부 다항식의 계수는 독립적으로 추정되어 질 수 있다. 퍼지 규칙이 할당된 공간의 지역적인 입력력 관계만을 독립적으로 고려하여 파라미터를 추정할 수 있다. 위에 설명한 바와 같이 지역적 학습을 통해 얻어진 퍼지 규칙의 후반부 구조인 다항식은 그 지역에서의 입력력 관계를 잘 나타내어 퍼지 모델의 장점인 시스템의 언어적 해석을 용이하게 할 수 있다 [3].

TSK 퍼지 추론 시스템의 일반적인 퍼지 규칙은 (1)과 같이 나타낼 수 있다.

$$R^j: \text{if } x_1 \text{ is } A_{1j} \wedge \dots \wedge x_m \text{ is } A_{mj}, \quad (1)$$

$$\text{then } y = f_j(x_1, x_2, \dots, x_m; \mathbf{a}_j)$$

본 논문에서는 TSK 퍼지 추론 시스템을 확장한 퍼지 결합 다항식 뉴럴 네트워크 기반으로 하여 패턴 분류 문제에 적용한다. Regression 문제에서 퍼지 규칙의 후반부 다항식의 계수를 추정하기 위하여 일반적으로 최소자승법을 사용한다. 패턴 분류 문제의 경우, 일반적으로 LDA를 기반으로 하여 각각의 퍼지 규칙들의 후반부를 독립적으로 추정 가능하도록 지역 학습을 이용한다. LDA의 지역 학습 형태는 WLDA이다.

### 2.1 Linear Discriminat Analysis (LDA)

Fisher에 의해 제안된 LDA 기법에서는 주어진 학습 데이터들이 (2)와 같은 선형식을 이용하여 원래 입력 공간의 차원보다 아주 작은 차원을 가지는 특성공간으로 투영된다[7].

$$y = \mathbf{w}^T \mathbf{x} \quad (2)$$

여기서  $y \in R$ 는 원래의 입력공간에 존재하는 데이터들을 특성공간으로 투영한 특성벡터이며,  $\mathbf{w} \in R^l$ 는 변환 벡터이다.

이  $l$ 은 입력변수의 개수를 의미한다.

LDA 기법은 (3)과 같이 정의된 목적 함수를 최대화 시키는 최적의  $\mathbf{w}$ 을 구하는 것이다.

$$J = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \quad (3)$$

여기서,  $\mathbf{S}_B$ 는 급간분산(between-class scatter) 행렬을 의미하며, (4)와 같이 구하며,  $\mathbf{S}_W$ 는 급내분산(within-class scatter) 행렬을 나타내며, (5)를 이용하여 구한다.

$$\mathbf{S}_B = (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T \quad (4-1)$$

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{x} \in D_i} \mathbf{x} \quad (4-2)$$

여기서,  $D_i$ 는  $i$ 번째 클래스에 속한 데이터들의 집합을 의미하며,  $n_i$ 는 집합  $D_i$ 에 속한 원소들의 개수를 의미한다.

$$\mathbf{S}_i = \sum_{\mathbf{x} \in D_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \quad (5-1)$$

$$\mathbf{S}_W = \mathbf{S}_1 + \mathbf{S}_2 \quad (5-2)$$

목적함수 (3)를 최대화 시키는  $\mathbf{w}$ 는 (6)과 같이 구할 수 있다.

$$\mathbf{S}_B \mathbf{w} = \lambda \mathbf{S}_W \mathbf{w} \quad (6-1)$$

$$\mathbf{w} = (\mathbf{S}_W)^{-1} (\mathbf{m}_1 - \mathbf{m}_2) \quad (6-2)$$

위에 언급한 전형적인 LDA는 입력공간의 모든 영역에 분포하는 모든 데이터를 대상으로 하여 (4)와 (5)를 이용하여 급간분산 행렬과 급내분산 행렬을 계산하였다.

### 2.2 Weighted Linear Discriminat Analysis (WLDA)

WLDA는 제안된 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기의 퍼지 규칙 후반부 구조인 다항식 뉴럴 네트워크의 다항식 계수를 추정하기 위하여 사용된다.

일반적인 LDA의 경우 모든 데이터의 하중치들이 동일한 경우에 적용가능하다. 그러나 퍼지 집합으로 입력공간이 분할된 경우 분할된 공간에 존재하는 데이터 패턴들의 하중은 동일하지 않게 된다. 각 데이터들의 하중 정보를 이용하여 LDA기법을 적용하기 위하여 WLDA 기법을 적용한다.

WLDA 기법도 일반적인 LDA와 동일한 형태의 목적함수인 (7)을 사용한다. 목적 함수 (7)은 1번째 퍼지 규칙에 관련된 다항식의 계수를 추정하기 위하여 정의된다.

$$J_l = \frac{\mathbf{w}_l^T \mathbf{S}_B^l \mathbf{w}_l}{\mathbf{w}_l^T \mathbf{S}_W^l \mathbf{w}_l} \quad (7)$$

퍼지 집합을 이용하여 공간 분할된 부 공간에서 데이터 패턴들이 가지고 있는 하중값(즉 퍼지 집합의 멤버십 함수의 값)을 가지고 있다. 이 하중값을 이용하여 퍼지 부공간

상에서 두 개의 클래스가 존재한다고 가정하면, 각 클래스의 평균 값 과 급간분산 행렬은 (8)과 같이 정의 된다.

$$S_B^i = (m_{i1} - m_{i2})(m_{i1} - m_{i2})^T \quad (8-1)$$

$$m_{ik} = \frac{\sum_{\mathbf{x} \in D_k} \mathbf{x} \cdot \mu_i(\mathbf{x})}{\sum_{\mathbf{x} \in D_k} \mu_i(\mathbf{x})} \quad (8-2)$$

여기서,  $D_k$ 는 k번째 클래스에 속한 데이터들의 집합을 의미하며,  $\mu_i$ 는 1번째 퍼지 집합의 멤버십 함수를 의미한다. 급내분산 행렬은 (9)와 같이 구할 수 있다.

$$S_{ik} = \sum_{\mathbf{x} \in D_k} (\mathbf{x} - m_{ik})\mu_i(\mathbf{x})(\mathbf{x} - m_{ik})^T \quad (9-1)$$

$$S_W^i = S_{i1} + S_{i2} \quad (9-2)$$

1번째 클러스터(즉, 1번째 퍼지 규칙)에서 목적함수 (7)를 최소화 시키는  $w_i$ 는 (10)과 같이 구할 수 있다.

$$S_B^i w_i = \lambda_i S_W^i w_i \quad (10-1)$$

$$w_i = (S_B^i)^{-1}(m_{i1} - m_{i2}) \quad (10-2)$$

### 3. 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기

제안된 퍼지 패턴 분류기는 독립적으로 학습되어지고 생성된 다수의 다항식 뉴럴 네트워크들이 선형 결합된 형태의 패턴 분류기이다. 독립적인 다수의 다항식 뉴럴 네트워크들의 결합은 (11)과 같은 퍼지 규칙에 의해 이루어진다.

$$R^j: \text{ if } \mathbf{x} \text{ is } C_j \text{ then } y = PNN_j(\mathbf{x}) \quad (11)$$

퍼지 규칙이 (11)과 같은 퍼지 규칙 기반 시스템의 최종 출력은 (12)와 같다.

$$\hat{y} = \frac{\sum_{i=1}^m \mu_i}{\sum_{k=1}^m \mu_k} \cdot PNN_i(\mathbf{x}) \quad (12)$$

여기서,  $\mu_k(\mathbf{x})$ 는 입력 데이터  $\mathbf{x}$ 가 퍼지 집합  $C_k$ 에 속하는 소속정도를 의미한다.

#### 3.1 다항식 뉴럴 네트워크

제안된 퍼지 규칙 기반 패턴 분류기의 후반부 구조로 사용되는 다항식 뉴럴 네트워크는 GMDH 알고리즘을 기반으로 하여 다변수 다차식 구조의 확장된 부분 표현식을 이용하여 연속적인 층과 노드를 생성하는 네트워크이다.

다항식 뉴럴 네트워크는 기존의 신경망 알고리즘에 비해 적은 노드수와 유연한 적응특성을 가지고 있다[5, 6]. 다항식 뉴럴 네트워크의 기본 노드인 확장된 다항식의 계수를 추정하기 위하여 지역 학습 알고리즘의 일종인 WLDA를 사용한다. 다항식 뉴럴 네트워크의 구조는 그림 1과 같다.

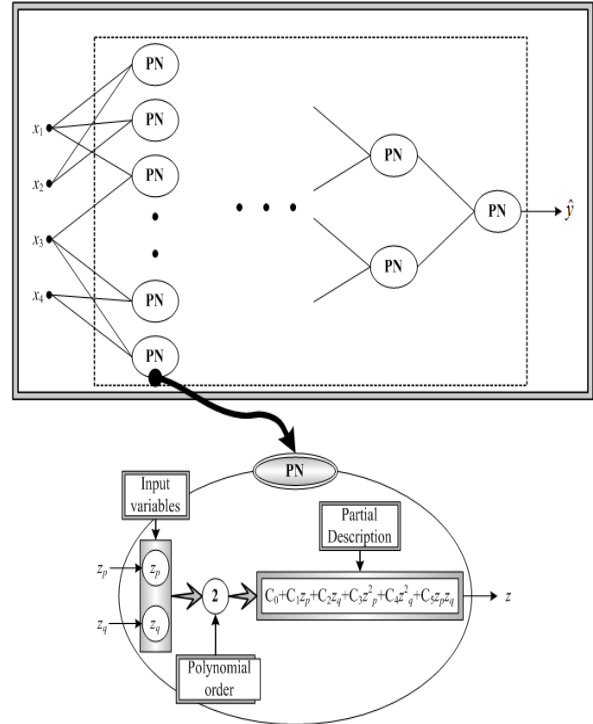


그림 1 다항식 뉴럴 네트워크의 구조  
Fig. 1 Structure of Polynomial Neural Networks

그림 1에서 PN은 다항식 뉴럴 네트워크의 기본 노드인 다항식 뉴런 (Polynomial Neuron; PN)을 의미한다. 여러 개의 노드들 중에서 미리 정의된 성능 지수를 만족하는 노드들이 선택되고 선택된 노드들은 다음단의 입력으로 사용된다.

이와 같은 과정을 되풀이 하게 되면, 복잡한 구조를 다항식을 만들게 된다. 다항식 뉴럴 네트워크의 기본 노드인 다항식 뉴런으로 사용되는 다항식의 형태는 표 1에 보인다.

표 1 다항식 뉴런의 다양한 형태

Table 1 Different forms of the regression polynomials used as polynomial neuron

No. of variables Type	1	2	3
0	$a$	N/A	N/A
1	$a_0 + a_1 \cdot x$	$a_0 + a_1 \cdot x_1 + a_2 \cdot x_2$	$a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3$
2	$a_0 + a_1 \cdot x + a_2 \cdot x^2$	$a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_1^2 + a_4 \cdot x_2^2 + a_5 \cdot x_1x_2$	$a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_1^2 + a_4 \cdot x_2^2 + a_5 \cdot x_3^2 + a_6 \cdot x_1x_2 + a_7 \cdot x_1x_3 + a_8 \cdot x_2x_3$
3		$a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_1x_2$	$a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_1x_2 + a_4 \cdot x_1x_3 + a_5 \cdot x_2x_3$

다항식 뉴럴 네트워크의 설계 방법은 다음과 같다.

#### Step 1. 시스템 입력변수의 결정 :

출력변수  $y$ 에 관계하는 n개의 시스템 입력변수를 결정한다. 이 시스템 입력변수를  $x_1, x_2, \dots, x_n$ 이라 한다.

**Step 2.** 데이터 분할 :

l개의 입력 출력 데이터  $(\mathbf{x}_j, y_j) = (x_{1j}, x_{2j}, \dots, x_{nj}, y_j)$ ,  $j=1, 2, \dots, l$  을 학습용 데이터와 테스트용 데이터로 분할한다.

**Step 3.** PN의 입력 변수 수 결정 :

PNN의 각 층에서 PN의 입력으로 선택될 수 있는 입력 변수의 수를 결정한다.

**Step 4.** PN의 입력 변수와 다항식 차수 결정 :

입력 변수의 수와 다항식의 차수의 종류를 이용하여 만들 수 있는 모든 조합을 이용하여 현재 층의 PN의 구조 (입력 변수, 다항식의 차수)를 결정한다. 현재 층에서 생성되어질 수 있는 중복되지 않는 입력변수와 다항식 차수를 가진 PN의 경우의 수는 식 (13)을 통해 구할 수 있다.

$$S = pCq \cdot T = \frac{p!}{(p-q)!q!} \cdot T \quad (13)$$

여기서, S는 중복되지 않는 PN의 경우 수, p는 현재 층의 입력으로 사용되는 전 층의 출력변수 수(현재 층이 1층이면 p는 시스템 입력변수의 수), q는 step 3에서 결정된 PN의 최대 입력 가능한 변수의 수, T는 다항식의 종류의 수이다 (본 논문에서는 T=4).

**Step 5.** LDA를 이용한 다항식 계수 추정 :

Step 4 에서 결정된 각 노드의 입력변수와 다항식의 차수를 이용하여 모든 노드의 구조를 결정하고, 식 (3)을 최대화 시키는 계수  $\mathbf{w}^*$ 을 식 (6-2)을 통해 구한다.

**Step 6.** 모델 테스트 및 노드의 선택 :

학습 데이터를 사용하여 뉴럴 다항식 뉴럴 네트워크의 분류 성능을 평가하고, 분류기의 성능을 평가하기 위해 각 PN의 각 클래스의 성능지수인 PI를 이용하여 평가한다. 즉, 다음 단의 입력으로 사용될 노드들은 (14) 최대화 시키는 노드들로 선택된다.

$$PI = \frac{((\mathbf{w}^*)^T \mathbf{m}_1 - (\mathbf{w}^*)^T \mathbf{m}_2)}{\sum_{\mathbf{x} \in D_1} ((\mathbf{w}^*)^T \mathbf{x} - (\mathbf{w}^*)^T \mathbf{m}_1)^2 + \sum_{\mathbf{x} \in D_2} ((\mathbf{w}^*)^T \mathbf{x} - (\mathbf{w}^*)^T \mathbf{m}_2)^2} \quad (14)$$

**Step 7.** PNN의 종료 판정 :

현재 i 층에서 얻어진 성능지수들 중에서 최대의  $PI_i^*$ 가 다음의 부등식을 만족하는 경우에 알고리즘을 종료한다.

$$PI_i^* \leq PI_{(i-1)}^* \quad (15)$$

단,  $PI_{(i-1)}^*$ 는 이전 층인 (i-1)층에서의 최대  $PI^*$ 이다.

**Step 8.** 다음 층의 새로운 입력변수 설정:

현재 층인 i층에서 보존된 노드의 출력들( $z_{i1}, z_{i2}, \dots, z_{iw}$ )은 다음 층인 (i+1)층의 새로운 입력들( $x_{(i+1)1}, x_{(i+1)2}, \dots, x_{(i+1)w}$ )로 식 (16)을 통해 구성된다. 원래 시스템 입력변수들이 추가되어 새로운 층의 입력변수로 정의된 후,  $i=i+1$ 로 변경하고

step 4로 간다.

$$x_{(i+1)1}^1 = z_{i1}, x_{(i+1)1}^2 = z_{i2}, \dots, x_{(i+1)1}^w = z_{iw} \quad (16)$$

여기서,  $z_i = \mathbf{w}^T \mathbf{x}_{i-1}$

**3.2 퍼지결합 다항식 뉴럴 네트워크 기반 패턴 분류기**

규칙 기반 패턴 분류기의 일종인 퍼지 규칙 기반 패턴 분류기의 성능을 향상시키기 위하여 퍼지 규칙의 후반부 구조로 다항식 뉴럴 네트워크를 적용한 퍼지 규칙 기반 패턴 분류기를 제안한다. 제안된 퍼지 규칙 기반 패턴 분류기의 전반부 구조는 입력의 다차원 문제를 해결하기 위하여 Radial Basis Function을 기반으로 한 퍼지 클러스터링 기법인 Fuzzy C-Means(FCM) 클러스터링 방법을 이용하여 정의한다.

**3.2.1 퍼지 규칙 전반부: 퍼지 클러스터링**

FCM 클러스터링은 n개의 벡터  $x_i (i=1, \dots, n)$  집합을 c개의 클러스터로 분할하고, 목적함수가 최소가 일 때 생성된 각 클러스터에서 중심 값을 찾는다. FCM 클러스터링의 목적함수는 식(17)와 같다.

$$\mathcal{J}(\boldsymbol{\mu}, \mathbf{v}) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m (\|\mathbf{x}_k - \mathbf{v}_i\|)^2 \quad (17)$$

목적함수를 최소화하는  $\boldsymbol{\mu}$ 와  $\mathbf{v}$ 는 식(6), (7)과 같이 얻을 수 있다.

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i\|}{\|\mathbf{x}_k - \mathbf{v}_j\|} \right)^{2/m-1}} \quad (18)$$

$$v_{ij} = \frac{\sum_{k=1}^N (u_{ik})^m x_{ki}}{\sum_{k=1}^N (u_{ik})^m} \quad (19)$$

**3.2.2 퍼지 규칙 후반부: 다항식 뉴럴 네트워크**

제안된 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기의 퍼지 규칙 후반부는 다항식 뉴럴 네트워크 이다. 앞에서 설명한 다항식 뉴럴 네트워크 설계 방법은 단독 네트워크에 대한 설계 방법이기 때문에 다수개의 다항식 뉴럴 네트워크를 구성하고 결합하기 위해서는 다항식 뉴럴 네트워크의 설계 방법이 수정되어야 한다.

다시 말하면, 앞에서 설명한 다항식 뉴럴 네트워크는 전체 입력공간을 모두 고려한 설계 방법인 반면, 제안된 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기의 다항식 뉴럴 네트워크는 퍼지 집합이 정의한 임의의 부 공간에서 다항식 뉴럴 네트워크의 설계가 이루어 져야 한다.

다음은 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기의 설계 방법을 나타낸다.

**Step 1.** 시스템 입력변수의 결정 :

출력변수  $y$ 에 관계하는  $n$ 개의 시스템 입력변수를 결정한다. 이 시스템 입력변수를  $x_1, x_2, \dots, x_n$  이라 한다.

**Step 2.** 데이터 분할 :

$l$ 개의 입출력 데이터  $(\mathbf{x}_j, y_j) = (x_{1j}, x_{2j}, \dots, x_{nj}, y_j)$ ,  $j=1, 2, \dots, l$  을 학습용 데이터와 테스트용 데이터로 분할한다.

**Step 3.** 퍼지 규칙의 수 결정 :

전체 입력 공간을 분할하는 퍼지 규칙의 수 ( $c$ )를 결정한다.

**Step 4.** 각 규칙의 후반부 구조인 다항식 뉴럴 네트워크를 설계 한다.

**Step 4.1.** 퍼지 규칙의 퍼지 집합의 중심 값 및 적합도 계산 :

퍼지 클러스터링 방법인 (18)과 (19)를 이용하여 퍼지 집합의 중심 값과 적합 도를 구한다.

**Step 4.2.** PN의 입력 변수 수 결정 :

PNN의 각 층에서 PN의 입력으로 선택될 수 있는 입력 변수의 수를 결정한다.

**Step 4.3.** 각 PN의 입력 변수와 다항식 차수 결정 :

입력 변수의 수와 다항식의 차수의 종류를 이용하여 만들 수 있는 모든 조합을 이용하여 현재 층의 PN의 구조 (입력 변수, 다항식의 차수)를 결정한다. 현재 층에서 생성되어질 수 있는 중복되지 않는 입력변수와 다항식 차수를 가진 PN의 경우의 수는 식 (13)을 통해 구할 수 있다.

**Step 4.4.** LDA를 이용한 다항식 계수 추정 :

Step 4.3 에서 결정된 각 노드의 입력변수와 다항식의 차수를 이용하여 모든 노드의 구조를 결정하고, 식 (7)을 최대화 시키는 계수  $w_i^*$ 을 식 (10-2)을 통해 구한다.

**Step 4.5.** 모델 테스트 및 노드의 선택:

학습 데이터를 사용하여 뉴럴 다항식 뉴럴 네트워크의 분류 성능을 평가하고, 분류기의 성능을 평가하기 위해 각 PN의 각 클래스의 성능지수인  $PI$ 를 이용하여 평가한다. 즉, 다음 단의 입력으로 사용될 노드들은 (20) 최대화 시키는 노드들로 선택된다.

$$PI = \frac{((w_i^*)^T m_{11} - (w_i^*)^T m_{12})^2}{\sum_{\mathbf{x} \in D_1} ((w_i^*)^T \mathbf{x} - (w_i^*)^T m_{11})^2 + \sum_{\mathbf{x} \in D_2} ((w_i^*)^T \mathbf{x} - (w_i^*)^T m_{12})^2} \quad (20)$$

**Step 4.6.** PNN의 종료 판정:

현재  $i$  층에서 얻어진 성능지수들 중에서 최대의  $PI_i^*$ 이 (15)을 만족하는 경우에 다항식 뉴럴 네트워크의 설계를 종료한다.

**Step 4.7.** 다음 층의 새로운 입력변수 설정 :

현재 층인  $i$ 층에서 보존된 노드의 출력들( $z_{i1}^j, z_{i2}^j, \dots, z_{iw}^j$ )

은 다음 층인 ( $i+1$ )층의 새로운 입력들( $x_{(i+1)}^1, x_{(i+1)}^2, \dots, x_{(i+1)}^w$ )로 식 (21)을 통해 구성된다. 원래 시스템 입력변수들이 추가되어 새로운 층의 입력변수로 정의 된 후,  $i=i+1$ 로 변경하고 Step 4.3 으로 간다.

$$x_{(i+1)}^1 = z_{i1}^j, x_{(i+1)}^2 = z_{i2}^j, \dots, x_{(i+1)}^w = z_{iw}^j \quad (21)$$

여기서,  $z_i = w_i^T x_{i-1}$ .

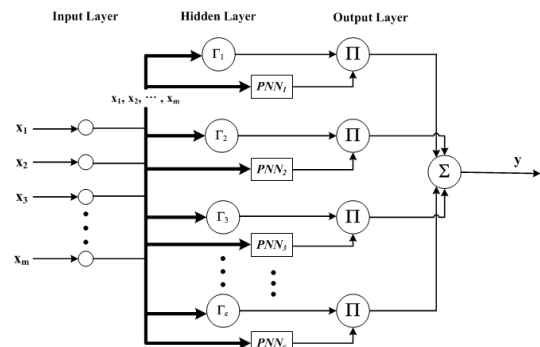
**Step 5.** step 4가 종료되었으면,  $l=l+1$ 로 변경하고 다음 퍼지 규칙의 후반부 구조를 결정하기 위하여 Step 4로 이동한다.

**Step 6.** 모든 퍼지 규칙에 대하여 후반부 구조 결정이 종료 되었으면 퍼지 결합 다항식 기반 패턴 분류기의 최종 출력을 (22)와 같이 구한다.

$$L(\mathbf{x}) = \begin{cases} 0, & \text{if } \hat{y}(\mathbf{x}) < 0 \\ 1, & \text{if } \hat{y}(\mathbf{x}) \geq 0 \end{cases} \quad (22)$$

$$\text{여기서, } \hat{y}(\mathbf{x}) = \frac{\sum_{k=1}^c \mu_k(\mathbf{x}) \cdot PNN_k(\mathbf{x})}{\sum_{k=1}^c \mu_k(\mathbf{x})}$$

그림 2는 퍼지 결합 다항식 뉴럴 네트워크의 구조를 나타낸다.



**그림 2** 퍼지 결합 다항식 뉴럴 네트워크의 구조  
**Fig. 2** The structure of fuzzy combined polynomial neural networks

**4. 시뮬레이션 및 결과 고찰**

본 논문에서 제안한 퍼지 결합 다항식 뉴럴 네트워크 기반으로 한 패턴 분류기의 성능을 평가하기 위하여 여러 개의 기계 학습 데이터들 사용하였다. 기계 학습 데이터 집합은 대표적인 기준 데이터 집합인 UCI 기계학습 데이터 저장소로부터 얻은 5가지의 데이터 집합들이다.

기준 데이터 집합을 이용하여 제안된 패턴 분류기의 성능과 특성을 기존 논문에서 이미 제안된 패턴 분류기의 성능과 비교, 평가 한다. 연구되어진 기존 패턴 분류기와 비교하

기 위하여, 전체 데이터를 학습 데이터와 테스트 데이터로 나누어 실험 한다.

데이터를 나누는 방법은 5 Fold Cross Validation 과 10 Fold Cross Validation 이며, 이는 참조할 논문들의 실험 조건과 동일한 조건에서 제안된 패턴 분류기와 기존의 패턴 분류기의 성능을 비교하기 위해서 이다.

표 2는 제안된 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기의 설계를 위해 미리 설정되어야 하는 파라미터들을 보인다.

**표 2** 제안된 패턴 분류기의 파라미터

**Table 2** Selected numeric values of the parameters of the proposed pattern classifier

Parameter	Value
Number of fuzzy rules (C)	2, 3, 5, 7
Number of input variables for each node	2, 3
Number of layers (L)	3, 5
Types of polynomials (T)	0, 1, 2, 3

제안된 패턴 분류기의 성능을 평가하기 위한 기계학습 데이터에 대한 개략적인 정보는 표 3에 열거 하였다.

**표 3** 실험에 사용된 기계학습 데이터

**Table 3** Machine Learning dataset used in the experiments

Dataset	Number of features	Number of Data	Number of Classes
Australian	42	690	2
Diabetes	8	768	2
German	24	1000	2
Liver	6	347	2
Sonar	60	208	2
Heart	13	270	2

표 4는 5-Fold Cross Validation 방법을 이용하여 데이터를 분할한 후 제안된 패턴 분류기와 다른 분류기와 데이터 패턴 분류성능을 비교한 것이다. 패턴 분류 성능은 오분류율로 비교하였다. 표 4에 보이는 바와 같이 제안된 패턴 분류기의 성능은 기존 LVQ 분류기 와 SVM에 비해 대체적으로 우수한 성능을 보임을 알 수 있다. 특히 Australian 데이터 경우에 특히 우수한 성능을 보였다.

표 5는 10 Fold Cross Validation에 의한 데이터 분할 후 실험 결과를 분석, 비교한 결과이다.

표 4, 5에서 비교된 기존 패턴 분류기는 WEKA라는 패키지를 이용하여 실험한 결과이다. 다양한 종류의 패턴 분류기와 패턴 분류 성능을 비교한 결과 대부분의 데이터에 대해 우수한 성능을 보임을 알 수 있다.

**표 4** 제안된 패턴 분류기와 기존 패턴 분류기 성능 비교(5-Fold Cross Validation)

**Table 4** Comparative analysis results of misclassification rates for proposed and conventional pattern classifiers (5-Fold Cross Validation)

Dataset	Proposed Classifier	LVQ1 [9]	LVQ2 [9]	LVQ3 [9]	SVM1 [9]	SVM2 [9]	SVM3 [9]
Australian	<b>15.51</b>	33.3	34.0	31.1	35.0	36.1	44.5
Diabetes	<b>24.87</b>	26.4	25.8	26.0	24.3	24.1	34.9
German	<b>27.3</b>	30.1	28.5	28.7	30.0	29.2	30.0
Heart	<b>30.25</b>	36.0	34.0	34.0	39.2	34.5	44.5
Liver	<b>32.17</b>	32.5	32.7	33.6	31.9	42.0	40.8

**표 5** 제안된 패턴 분류기와 기존 패턴 분류기 성능 비교 (10 Fold Cross Validation)

**Table 5** Comparative analysis results of misclassification rates for proposed and conventional pattern classifiers (10 Fold Cross Validation)

	Australian	Liver	German	Diabetes	Sonar	Heart
1NN	18.43	37.78	28.12	29.38	13.83	N/A
kNN (k=3)	15.04	37.52	27.79	26.14	16.24	N/A
J48	14.43	34.16	28.75	25.51	26.39	N/A
PART	15.55	34.75	29.46	26.55	22.6	N/A
Bayes networks	22.14	45.11	24.84	24.25	32.29	N/A
SMO	15.12	42.02	24.91	23.2	23.4	N/A
RBFNN	20.45	34.94	26.42	25.96	27.38	N/A
LVQ3[9]	31.1	33.6	28.7	26.0	N/A	34.0
SVM[9]	35.0	31.9	30.0	24.3	N/A	39.2
Proposed Classifier	<b>14.82</b>	<b>31.10</b>	<b>25.6</b>	<b>23.57</b>	<b>30.79</b>	<b>17.78</b>

## 5. 결 론

본 논문에서는 퍼지 규칙 기반 패턴 분류기의 일종인 퍼지 결합 다항식 뉴럴 네트워크 기반 패턴 분류기를 제안하였다. 제안된 퍼지 결합 다항식 뉴럴 네트워크는 기존 퍼지 규칙 기반 시스템인 퍼지 모델의 후반부 구조를 기존의 일차 선형 다항식이 아닌 다항식 뉴럴 네트워크를 채택한 구조이다.

유연한 구조 선택 능력과 단순한 구조를 단계적으로 증척시키는 구조인 다항식 뉴럴 네트워크는 복잡한 구조임에도 불구하고, 다항식의 계수 추정 부분에서 우수한 성능을 보인다. 또한 입력 공간 분할을 위해서는 다차원 공간 문제를 해결하기 위하여 FCM 클러스터링 방법을 사용하였다. 제안된 퍼지 결합 다항식 뉴럴 네트워크의 패턴 분류 성능은 시뮬레이션 결과를 통해 알 수 있듯이 기존 패턴 분류 알고리즘들에 비해 우수함을 알 수 있다.

퍼지 규칙의 전반부인 퍼지 집합에 의해 정의된 지역에서, 다항식 뉴럴 네트워크를 이용하여 패턴 분류를 위한 다

양한 형태의 Discriminant 함수를 구현함으로써 우수한 패턴 분류 성능을 얻는다. 입력 공간을 퍼지 분할하는 퍼지 집합과 분할된 퍼지 공간에 할당된 다항식 뉴럴 네트워크를 최적화시킴으로써 좀 더 개선된 패턴 분류 성능을 얻을 수 있을 것으로 판단된다.

**감사의 글**

본 논문은 2013학년도 원광대학교의 교비지원에 의해서 수행 됨

**References**

[1] G. Pajares, M. Guijarro, A. Ribeiro, "A Hopfield Neural Network for combining classifiers applied to textured images," Neural Networks, Vol.23, pp. 144-153, 2010.

[2] S.-K. Oh and Pedrycz, W, "The Design of Self-Organizing Polynomial Neural Networks," Information Sciences, Vol. 141, pp. 237-258, 2002.

[3] John Y., Liang W., and Charles W. G., "Improving the Interpretability of TSK Fuzzy Models by Combining Global Learning and Local Learning", IEEE Trans. on Fuzzy Systems, vol. 6, no. 4, 1998.

[4] Pedrycz W. "An identification algorithm in fuzzy relation system," Fuzzy Sets and Systems, vol 13, pp. 153-167, 1984.

[5] Oh S. K., Pedrycz W., and Park B. J., "Polynomial Neural Networks Architecture: Analysis and Design," Computers and Electrical Engineering, vol. 29, no. 6, pp. 703-725, 2003.

[6] Ivahnenko A. G., "Polynomial theory of complex systems," IEEE transaction on Systems, Man and Cybernetics, vol. SMC-12, pp. 364-378, 1971.

[7] Na J. H., Park M. S., and Choi J. Y., "Linear boundary discriminant analysis," Pattern Recognition, vol. 43, pp. 929 - 936, 2010.

[8] Pedrycz W., "Conditional fuzzy clustering in the design of radial basis function neural networks," IEEE Transaction on Neural Networks, vol. 9, no.4, pp. 601-612, 1998.

[9] Parades R, Vidal E., "Learning prototypes and distance: A prototype reduction technique based on nearest neighbor error minimization," Pattern Recognition, vol. 39, pp. 180-188, 2006.

[10] Seok-Beom Roh, Kyung-Won Jang, Kuk-Yeon Hwang and Tae-Chon Ahn, "The design of Pattern Classifiers based on Fuzzy combined Polynomial Neural Networks," Proceedings of KIEE Summer Conference 2013, pp.1611-1612, 2013.

**저 자 소 개**



**노 석 범 (盧石範)**

1994년 원광대학교 제어계측공학과 졸업, 1996년 동 대학원 컴퓨터공학과 석사, 2006년 동 대학원 제어계측공학과 박사, 2007년~현재 원광대학교 공과대학 전자융합공학과 연구교수  
관심분야 : 퍼지 이론, 신경 회로망, Bio-inspired optimization algorithm, Pattern Recognition



**장 경 원 (張京源)**

2001년 원광대학교 전기공학부 졸업, 2003년 동 대학원 제어계측공학과 석사, 2008년 동 대학원 제어계측공학과 박사, 2008년~현재 원광대학교 공과대학 전자융합공학과 겸임교수  
관심분야 : 퍼지 이론, 신경 회로망, Bio-inspired optimization algorithm, Pattern Recognition



**안 태 천 (安泰天)**

1978년 연세대학교 전기공학과 졸업 1980년 동 대학원 전기공학과 석사, 1986년 동 대학원 전기공학과 박사, 1981년~현재 원광대학교 공과대학 전자융합공학과 교수.  
관심분야 : 퍼지 이론 및 응용, 신경망, Bio-inspired optimization algorithm, Intelligent Pattern Recognition, Data mining, 태양광.