

# 퍼지 kNN과 Conditional FCM을 이용한 퍼지 RBF의 설계

논 문

58-6-25

## Design of Radial Basis Function with the Aid of Fuzzy KNN and Conditional FCM

노 석 범\* · 오 성 권†

(Roh, Seok-Beon · Oh, Sung-Kwun)

**Abstract** - The performance of Radial Basis Function Neural Networks depends on setting up the Radial Basis Functions over the input space which are the important design procedure of Radial Basis Function Neural Networks. The existing method to initialize the location of the radial basis functions over the input space is to use the conditional fuzzy C-means clustering. However, the researchers which are interested in the conditional fuzzy C-means clustering cannot get as good modeling performance as they expect because the conditional fuzzy C-means clustering cannot project the information which is extracted over the output space into the input space. To compensate the above mentioned drawback of the conditional fuzzy C-means clustering, we apply a fuzzy K-nearest neighbors approach to project the auxiliary information defined over the output space into the input space without loss of the information.

**Key Words** : Radial Basis Function Neural Networks(RBFNN), Conditional Fuzzy C-Means(CFCM), Fuzzy k-Nearest Neighbors(FKNN), Auxiliary Information

### 1. 서 론

Radial Basis Function (RBF) 뉴럴 네트워크는 수많은 연구자들에 의해 Regression 문제와 Classification 문제와 같은 다양한 분야에 적용하기 위한 연구가 진행되어지고 있다 [2-4]. 특히 RBF 뉴럴 네트워크의 초기 모델인 제안된 후[1], RBF 뉴럴 네트워크의 초기 구조를 변형한 새로운 모델과 학습 방법이 제안되었다[5-9]. RBF 뉴럴 네트워크의 중요한 특징은 RBF의 파라미터들이 출력 층의 연결 하중과 별개로 결정 된다는 것이다. 일반적인 뉴럴 네트워크의 오류 역전과 알고리즘과 같은 비선형 최적화 방법과 비교해서 RBF 뉴럴 네트워크에 적용된 학습 방법이 매우 빠르며, 실제 문제에 적용할 경우 지역 최소값(local minima)에 수렴을 회피 할 수 있다고 알려져 있다[10]. RBF 뉴럴 네트워크의 기본 모델은 RBF를 은닉 층으로 가지며, 출력 층의 노드는 은닉 층에 존재하는 노드(RBF)들의 선형 결합으로 구성된다. 제안된 RBF 뉴럴 네트워크의 기본 모델을 확장 발전시키기 위한 측면으로 크게 두 가지가 고려되어진다. 첫 번째는 은닉 층의 최적화로써, RBF의 일반적인 구조는 gaussian 함수이지만 이를 변형한 다양한 RBF 형태가 연구되어지고 있으며, RBF의 파라미터(RBF의 위치)를 최적화하기 위한 방법들이 제안되었다. RBF의 위치를 결정하는 방법으로 K-means 와 fuzzy C-means와 같은 클러스터링 기법들이 사용되어진다[11-13]. 두 번째는 출력 층을 최적화 하는 방법이다. RBF 뉴럴 네트

워크의 은닉 층과 출력 층 사이의 연결 하중을 추정하기 위해서 일반적으로 최소 자승법을 사용하지만, Gradient Decent 방법과 Expectation Maximization (EM) 기법들도 제안되었다 [14-16]. RBF 뉴럴 네트워크의 초기 모델에서 RBF의 구조는 gaussian 함수 형태를 가지고 있으며, gradient decent 방법에 의한 학습을 통해 gaussian 함수의 중심 값과 분산을 결정한다. 초기 RBF 뉴럴 네트워크와 달리 입력 공간상에서의 클러스터링 기법을 통해 데이터 분포를 분석한 정보를 RBF의 구조와 파라미터 결정에 이용하는 네트워크가 제안되었다. RBF의 구조 및 파라미터 결정에 필요한 정보를 얻기 위한 다양한 클러스터링 방법 들 중에서 가장 적합한 클러스터링 기법으로 Bezdeck이 제안한 Fuzzy C-Means (FCM) clustering 기법이 많이 알려져 있다. 그러나 FCM 클러스터링 기법과 같은 unsupervised 클러스터링 기법은 오직 입력 공간상에 분포된 입력 변수의 패턴만을 해석하기 때문에 regression 문제에 적용할 경우 입 · 출력 변수의 관계를 해석하기에는 한계가 있다고 알려져 있다.

위에 언급한 unsupervised 클러스터링 기법을 regression 문제에 적용할 경우의 단점을 해결하기 위하여, 입 · 출력 변수의 관계를 고려한 클러스터링 기법들이 제안되었다. 입 · 출력 변수의 관계를 이용한 클러스터링 기법으로 Pedrycz가 제안한 conditional Fuzzy C-Means (c-FCM) 클러스터링 기법이 대표적이다. Pedrycz에 의해 제안된 c-FCM 기법은 출력 공간상에서 출력변수의 패턴 분포를 해석하고, 얻어진 정보를 참고하여 입력 공간상에서 입력변수의 패턴 분포를 해석하는 방법이다. c-FCM 기법을 적용한 경우 출력 공간상에 정의된 부가적 정보와 이에 상응하는 입력공간상에서 각 cluster의 prototype (중심 값) 과 activation level을 얻을 수 있다. 그러나 c-FCM 기법은 데이터 패턴들의 입 · 출력 쌍들

† 교신저자, 정회원 : 수원대학 전기공학과 정교수

E-mail : ohsk@suwon.ac.kr

\* 정 회 원 : 원광대학 제어계측 공학과 · 공박

접수일자 : 2008년 11월 19일

최종완료 : 2009년 3월 30일

이 알려져 있는 학습 데이터 집합에서는 출력 공간상에서의 부가적 정보를 얻을 수 있지만, 출력 데이터에 대한 정보가 전혀 없는 테스트 데이터 집합에서는 출력 공간상의 부가적 정보를 얻을 수 없다는 단점이 있다.

본 논문에서는 앞서 언급한 c-FCM의 단점을 보완하기 위하여 Fuzzy K-Nearest Neighbors (FKNN) 기법을 사용하여 출력공간상의 부가적인 정보를 추정하고, 추정된 정보를 이용하여 각 cluster에 새로운 데이터가 속하는 소속정도를 구하게 된다.

### 2. Radial Basis Function 뉴럴 네트워크

RBF 뉴럴 네트워크의 구조는 그림 1과 같다. 각각의 은닉층 노드는 입력 변수를 입력으로 하는 receptive field(다시 말하면, kernel function)의 값을 출력으로 가진다.

전형적인 RBF 뉴럴 네트워크는 식 (1)과 같은 gaussian 함수를 receptive field로 정의해서 사용한다.

$$\Phi_i(x) = \exp\left(-\frac{\|x - v_i\|^2}{2\sigma_i^2}\right) \quad (1)$$

여기서,  $v_i$ 와  $\sigma_i^2$ 는 각각  $i$ 번째 receptive field의 중심 값과 분산을 의미한다. 각각의 receptive field는 입력 변수 값이 해당 receptive field의 중심 값과 가까우면 가까울수록 1에 근접한 출력 값을 발생 시킨다. 입력변수의 값과 receptive field의 중심 값 사이의 거리가 멀어지면 멀어질수록 receptive field의 출력 값은 단조 감소한다.

식(1)과 같이 정의된 gaussian receptive field는 제안된 함수로 변형되어 사용되어진다. receptive field의 가중 선형 결합인 RBF 뉴럴 네트워크의 출력은 식(2)과 같다.

$$y(x) = \sum_{i=1}^c w_i \Phi_i(x) \quad (2)$$

여기서,  $w_i$ 는 은닉층의  $i$ 번째 노드와 출력 노드사이의 연결하중을 의미한다. Gaussian 함수를 RBF로 가진 전형적 RBF 뉴럴 네트워크의 구조는 앞서 설명하였다. 전형적인 RBF 뉴럴 네트워크와 달리 FCM 기반 RBF 뉴럴 네트워크는 식(1)과 같은 gaussian 함수를 receptive field로 사용하지 않고, FCM 기법을 통해 얻어진 중심값과 함수 형태를 receptive field로 사용한다.

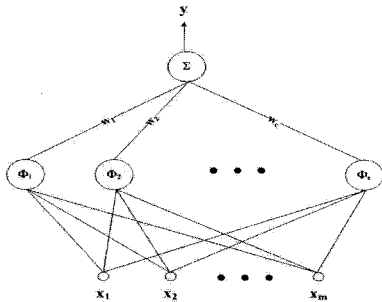


그림 1 RBF 뉴럴 네트워크의 구조  
Fig. 1 The topology of RBF Neural Networks

네트워크 구조 측면에서 전형적인 RBF 뉴럴 네트워크와 FCM 기반 RBF 뉴럴 네트워크는 동일한 구조를 공유한다. FCM 기법을 통해 구해진 receptive field의 함수는 식(3)과

같다.

$$\Phi_i(x_k) = \frac{1}{\sum_{j=1}^c \left( \frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{p-1}}} \quad (3)$$

### 3. Conditional Fuzzy Clustering 기법

RBF 뉴럴 네트워크를 설계하기 위하여 RBF들의 위치와 모양을 정의 하는 것이 RBF 뉴럴 네트워크의 모델링 성능을 결정하는 중요한 요소가 된다. RBF의 위치와 모양을 정의하기 위해 클러스터링 기법을 사용하는 것은 잘 알려진 사실이다. 기존의 클러스터링 방법은 입력공간상에서 분포된 데이터들의 분포를 분석하여 이해하는 것에 중점을 두고 있다. 클러스터의 할당은 미리 정의된 목적함수를 최소화하는 관점에서 행하여진다. 이러한 목적함수의 최소화 과정에서는 어떠한 교사신호도 사용되어지지 않는다. 이러한 관점에서 클러스터의 할당을 위한 unsupervised learning은 출력공간상에서 얻을 수 있는 어떤 부가적인 정보도 사용하지 않는다고 할 수 있다. RBF 뉴럴 네트워크를 설계한다는 측면에서는 출력공간상에서 얻을 수 있는 부가적인 정보를 사용하는 것이 RBF 뉴럴 네트워크의 성능 개선에 유리하다고 할 수 있다. 그러나 기존의 FCM 클러스터링 기법은 전적으로 unsupervised learning에 기반을 둔 방법이어서 앞서 언급한 출력공간상의 부가적인 정보를 전혀 사용할 수 없다는 단점이 있다. 이러한 단점을 극복하기 위한 방법으로 c-FCM 클러스터링 기법이 제안 되었다. 입력패턴들의 집합 X를 다음과 같이 정의한다.

$$X = \{x_1, x_2, \dots, x_n\}, x_k \in R^m \quad (4)$$

여기서 m은 입력공간의 차원을 의미하며, n은 데이터 패턴의 수를 의미한다. 입력패턴들의 집합 X에 대응하는 출력패턴들의 집합 Y는 (5)와같이 정의한다.

$$Y = \{y_1, y_2, \dots, y_n\}, y_k \in R \quad (5)$$

집합 X의 각각의 요소는 집합 Y의 요소들과 대응된다.

Conditional 클러스터링 기법에서 데이터 패턴  $x_k$ 는 출력공간상에서 정의된 임의의 언어적 변수에 의해 얻어진 부가적인 정보과 입력공간상에서의 데이터 패턴들의 분포를 고려하여 분류되어진다. 출력 공간상에서 얻어진 부가적인 정보인 Context와 얻어진 부가적인 정보를 이용하여 입력공간에서 입력 데이터 패턴을 분석하여 얻은 클러스터와의 관계는 그림 2와 같다.

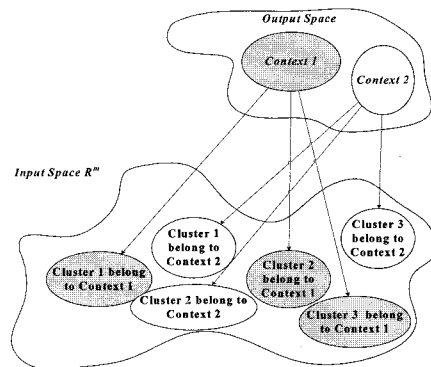


그림 2 Context 기반 클러스터링  
Fig. 2 Context based Clustering

그림 2에 보인 것과 같이 출력 공간에서 출력변수의 데이터 패턴들의 분포를 분석하여 Context를 정의한다. 출력공간상에서 정의된 Context는 언어적 변수인 퍼지 집합 B로 표현되며 (6)과 같이 정의된다.

$$B : R \rightarrow [0, 1] \quad (6)$$

출력패턴  $y_k$ 가 퍼지 집합 B에 소속되는 소속정도를 나타내는  $f_k$ 는 (7)와 같이 정의된다.  $f_k$ 는 부가적 정보인 Context의 멤버십 함수를 의미한다.

$$f_k = B(y_k), \quad 1 \leq k \leq n \quad (7)$$

c-FCM 기법을 위한 목적함수는 일반적인 FCM 알고리즘의 목적함수와 동일하며 (8)와 같이 정의된다.

$$J_c = \sum_{j=1}^r \sum_{k=1}^n (\mu_{jk}^l)^p \cdot \|x_k - v_j^l\|^2, \quad 1 \leq l \leq C \quad (8)$$

여기서  $J_c$ 는 1번째 context (출력공간상에 정의된 퍼지 집합)에서 정의된 목적함수를 의미하며,  $\mu_{jk}^l$ 은 1번째 context에서 정의된 소속정도를 나타내는 행렬의 요소를 의미한다.  $v_j^l$ 은 1번째 context에서 얻어진 클러스터들의 prototype을 의미한다. C와 r은 각각 출력 공간상에서 정의된 context의 개수와 각 context와 연관된 입력공간상에 정의된 클러스터의 수를 의미한다. 기존의 FCM과 동일한 목적함수를 사용하는 c-FCM의 제한 조건은 FCM의 제한 조건과 다르다. FCM의 제한 조건은 (9)과 같이 정의되지만 c-FCM의 제한 조건은 (10)과 같이 정의된다.

$$\sum_{j=1}^r u_{jk} = 1, \quad 1 \leq k \leq n \quad (9)$$

$$\sum_{j=1}^r u_{jk}^l = f_k^l, \quad 1 \leq k \leq n \quad (10-1)$$

$$f_k^l = B_l(y_k) \quad (10-2)$$

여기서,  $B_l$ 은 출력공간상에서 정의된 1번째 언어적 변수를 의미하며 퍼지 집합으로 정의된다.

제한 조건 (10)을 만족하며 목적함수 (8)을 최소화 시키는 최적화 문제는 (11)과 같이 정의한다.

$$\min_{U, v} J_c \text{ subject to } \sum_{j=1}^r u_{jk}^l = f_k^l \quad (11)$$

최적화 문제 (11)을 최소화하기 위해 반복적 방법을 이용하여 얻어진 partition matrix  $U = [u_{ij}^l]$ 와 각 클러스터의 prototype은 (12)와 (13)과 같이 구할 수 있다.

$$\phi_i^l(x_k) = u_{ik}^l = \frac{f_k^l}{\sum_{j=1}^C \left( \frac{\|x_k - v_j^l\|}{\|x_k - v_i^l\|} \right)^{2/(p-1)}} \quad (12)$$

여기서,  $\phi_i^l$ 은 1번째 context와 관련된 i번째 클러스터의 멤버십 함수를 의미한다.

$$v_i^l = \frac{\sum_{k=1}^n (u_{ik}^l)^p \cdot x_k}{\sum_{k=1}^n (u_{ik}^l)^p} \quad (13)$$

학습 데이터 집합에 c-FCM을 적용하여 얻어진 식(12)와 (13)은 RBF 뉴럴 네트워크의 RBF를 정의한다. 테스트 데이터 집합에 (12)와 (13)을 적용하여 RBF의 출력 값을 구하고자 한다면, 식 (12)식의 분자인  $f_k^l$ 을 알아야 한다. 그러나 (10-2)에 보인 것과 같이  $f_k^l$  값을 알기 위해서는 출력변수의

값( $y_k$ )을 알아야 하지만, 테스트 데이터 집합의 경우는 출력 변수 값이 알려져 있지 않다고 가정하여 사용하기 때문에  $f_k^l$ 을 구할 수 없게 된다. 그러므로 지금까지 c-FCM을 이용하여 RBF를 정의하는 경우에는 (12)를 그대로 사용하지 못하고 (14)와 같이 변경하여 사용하였다.

$$u_{ik}^l = \frac{1}{\sum_{j=1}^r \left( \frac{\|x_k - v_j^l\|}{\|x_k - v_i^l\|} \right)^{2/(p-1)}} \quad (14)$$

(14)와 기존의 FCM에 의해 정의된 partition matrix를 정의한 식 (3)을 비교하면 동일하다는 것을 알 수 있다. (14)와 (3)이 동일하다는 의미는 출력공간상에서 추출된 부가적 정보  $B_l$ 이 각 클러스터의 prototype  $v_j^l$ 을 정의하기 위해서만 이용되고 RBF 뉴럴 네트워크의 partition matrix를 구하기 위한 식 (14)에는 이용되지 않게 되어 부가적 정보인 context를 최대한 활용 할 수 없게 된다는 것을 의미한다.

그러므로 출력공간에서 정의된 부가적인 정보  $B_l$ 를 RBF 뉴럴 네트워크의 partition matrix에도 반영하기 위해서는 새롭게 주어진 입력 패턴  $x_{new}$ 에 대응하는 출력공간상에서 정의된 부가적 정보  $f_{new}^l$ 를 추정할 필요성이 대두되었다.

본 연구에서는 부가적 정보  $f_{new}^l$ 추정하기 위한 방법으로 Fuzzy K-Nearest Neighbors (FKNN) 기법을 이용하였다. FKNN 기법을 이용하여 추정된  $f_{new}^l$ 을 식(12)에 적용하여 식 (12)를 (14)변경할 필요 없이 RBF를 정의 할 수 있다.

#### 4. Fuzzy K-Nearest Neighbors 기법을 이용한 부가적 정보 추정

앞에서 언급한 바와 같이 c-FCM 클러스터링 기법을 수행하기 위해서는 출력공간상에서 부가적 정보를 정의하는 퍼지 집합  $B_l$ 을 정의하여야 한다.

##### 4.1 출력 공간상에서 부가적 정보인 context 추출

출력공간상에서 부가적 정보 추출 방법을 출력 공간상에서 정보 입자화라 한다. 본 논문에서는 출력 공간상에서 출력변수들의 데이터 분포를 분석하여 부가적 정보를 추출하고 추출된 정보를 이용하여 퍼지 집합을 정의하기 위하여 균등 분할 기법과 K-means 클러스터링 기법을 적용하였다.

###### 4.1.1 균등 분할

출력공간상에서 데이터의 분포를 분석하기 위하여 학습 데이터 패턴들의 출력변수의 최솟값과 최댓값을 사이에 임의의 개수의 context들을 정의하는 분할 방법이다.

균등 분할 방법은 출력공간의 최댓값과 최솟값사이의 공간을 미리 정의된 개수의 공간으로 균등 분할하는 방법이다. 각 context의 정점은 (15)와 같이 구한다.

$$v_i = \frac{(y_{max} - y_{min})}{(C-1)} \cdot (i-1) + y_{min} \quad (15)$$

여기서, C는 context의 개수를 의미한다.

###### 4.1.2 K-means 클러스터링 기법을 이용한 공간분할

일반적으로 균등분할을 이용한 공간분할은 그 공간에 분포된 데이터 패턴들의 분포 패턴을 적절하게 반영하지 못하는 경향이 있는 것으로 알려져 있다. 이러한 특성을 개선하기 위한 방법으로 클러스터링 방법들 중 하나인 K-means 클러스터링 기법을 이용하여 출력공간상의 데이터 패턴들의 분포

를 파악하여 공간을 분할한다. 출력 공간을 균등 분할하는 분할방법과 달리 K-means 클러스터링을 이용한 공간 분할 방법으로 분할된 context들의 정점들은 불규칙하게 분포하는 특성을 보인다. 균등 분할 방식과 K-means 클러스터링 방식을 비교해 보면 균등 분할을 통해 얻어진 context들의 정점들이 출력공간에 균등하게 분포해 있는 반면, K-means 클러스터링에 의해 얻어진 정점들은 불규칙하게 분포되어지는 것을 알 수 있다. 본 논문에서는 Gaussian type 멤버십 함수는 사용하지 않고 삼각형 멤버십 함수를 사용한다.  $f_j$ 는  $i$ 번째 context에  $j$ 번째 출력데이터( $y_j$ )가 속한 소속정도를 의미하며, 이 소속정도는 c-FCM 클러스터링의 제한 조건으로 쓰인다(식 (10-1) 참조).

**4.2 FKNN 기법을 이용한 부가적 정보 입자의 추정**

앞서 설명한 바와 같이 c-FCM 클러스터링의 단점은 부가적 정보 입자인  $f_k$ 를 새롭게 주어지는 데이터 패턴에 대해서는 알 수 없을 경우 일반적인 FCM 클러스터링과 비슷한 결과들 도출한 다는 것이다. 이러한 단점을 극복하고 c-FCM을 이용하여 RBF 뉴럴 네트워크의 Radial Basis Function (RBF)를 정의하기 위하여, classification 기법의 일종인 FKNN 기법을 적용하여 부가적 정보 입자인  $f_k$ 를 추정하는 방법을 제안한다.

**4.2.1 Crisp K-Nearest Neighbors 기법**

대표적인 non-parametric 알고리즘인 K-Nearest Neighbors(KNN) 기법은 다양한 classification 문제에 적용되어져왔다. 일반적으로, KNN 기법은 알려져 있지 않은 데이터 패턴이 주어질 경우, 주어진 데이터 패턴과 가장 가까운 위치에 있는 k개의 데이터 패턴을 선정하고, 선정된 데이터 패턴의 출력 패턴을 비교하여 새롭게 주어진 데이터 패턴의 출력 값을 추정한다. Crisp K-Nearest Neighbors 기법을 설명하기 위하여 입력 패턴 집합 X와 입력 패턴에 대응하는 출력 패턴의 집합 Y를 (16)과 (17)와 같이 정의한다. Classification 문제에서 출력 패턴의 집합은 연속적인 값이 아닌 이산적인 값을 가진다.

$$X = \{x_1, x_2, \dots, x_n\}, x_k \in R^m \quad (16)$$

$$Y = \{y_1, y_2, \dots, y_n\}, y_k \in N \quad (17)$$

여기서, N은 자연수의 집합을 의미하며, n은 데이터 패턴의 수를 의미한다. 학습 데이터의 집합이 (16)와 같이 정의되어졌다고 가정하고, 분류해야할 query instance  $x_q$ 가 주어지면, 학습 데이터 집합 X의 요소들 중에서  $x_q$ 와 가까운 거리에 있는 k개의 데이터 패턴을 찾는다. 이때 데이터 패턴들 간의 거리는 (18)과 같이 정의된 weighted Euclidean distance를 이용하여 계산한다.

$$d(a, b) = \sqrt{\sum_{i=1}^m \frac{(a_i - b_i)^2}{\sigma_i^2}} \quad (18)$$

여기서,  $\sigma_i^2$ 은 1번째 입력변수의 분산을 의미한다. (18)에 의해 얻어진 query instance와 학습 데이터 집합의 요소들 사이의 거리 정보를 이용하여 거리가 가까운 K개의 학습 데이터들은 찾는다. 집합 L은 주어진 query instance와 거리가 가까운 데이터 패턴의 인덱스의 집합을 의미한다.

$$L = \{j: x_j \text{ is one of the } k \text{ nearest instances to query point } q\} \quad (19)$$

근접 위치에 있는 K개의 학습 데이터들의 출력 label을 이용하여 주어진 query instance의 출력 label을 (20)과 같이 구한다.

$$\hat{y}(x_q) = \arg \max_c \sum_{j=1}^k f(y(x_{L(j)}), c) \quad (20)$$

$$\text{여기서, } f(y(x_{L(j)}), c) = \begin{cases} 1, & \text{if } y(x_{L(j)}) = c \\ 0, & \text{if } y(x_{L(j)}) \neq c \end{cases}$$

**4.2.2 Fuzzy K-Nearest Neighbors 기법**

Fuzzy K-Nearest Neighbors 기법은 기존의 Crisp K-Nearest Neighbors 기법에 퍼지 집합의 개념을 적용한 기법으로 주어진 query instance와 학습 데이터 패턴들 간의 유사성(similarity)의 정의에 기반을 두고 있다.

데이터 패턴들 간의 유사성은 (21)과 같이 정의한다.

$$S(x_{L(i)}, q) = \begin{cases} 1, & \text{if } x_{L(i)} = q \\ \frac{1}{\sum_{j=1}^K \left( \frac{\|x_{L(i)} - q\|}{\|x_{L(j)} - q\|} \right)^{\frac{2}{p-1}}}, & \text{if } x_{L(i)} \neq q \end{cases} \quad (21)$$

여기서, S는 데이터  $x_{L(i)}$ 와 query instance  $q$ 와의 유사성을 의미하며, K는 Nearest neighbor들의 수, p는 퍼지화 계수(fuzzification coefficient)를 의미한다. (21)를 통해 얻어진 유사성  $S(x_{L(i)}, q)$ 는 다음 조건을 만족한다.

$$0 < S(x_{L(i)}, q) \leq 1 \quad (22)$$

근접 위치에 있는 K개의 학습 데이터들의 출력 label을 이용하여 fuzzy K-Nearest Neighbor 기법을 적용하여 주어진 query instance의 출력 label은 (23)과 같이 구한다.

$$\hat{y}(q) = \sum_{i=1}^K S(x_{L(i)}, q) \cdot y(x_{L(i)}) = \sum_{i=1}^K \left( \frac{y(x_{L(i)})}{\sum_{j=1}^K \left( \frac{\|x_{L(i)} - q\|}{\|x_{L(j)} - q\|} \right)^{\frac{2}{p-1}}} \right) \quad (23)$$

여기서,  $\hat{y}(q)$ 는 query instance와 연관된 fuzzy K-Nearest Neighbor 기법을 이용하여 추정된 출력 label을 의미하며,  $y(x)$ 는 데이터 패턴  $x$ 와 연관된 알려진 출력 label을 의미한다.

**4.2.3 FKNN 과 c-FCM 기반 RBF 뉴럴 네트워크의 설계**

Pedrycz가 제안한 c-FCM 클러스터링 기법을 이용하여 RBF 뉴럴 네트워크의 receptive field를 정의하기 위하여 출력 공간상에 정의된 부가적 정보인 context의 정의가 필요하다. 출력변수의 값이 알려져 있지 않은 테스트 데이터 집합의 각각의 context에 대한 소속정도를 계산하기 위하여 위에 설명한 FKNN 기법을 이용하여 테스트 데이터 집합의 부가적 정보를 추정한다. 제안된 RBF 뉴럴 네트워크의 구조는 그림 3과 같다. 그림 3에 보인 바와 같이, 제안된 RBF 뉴럴 네트워크의 출력은 은닉층의 노드인 RBF의 출력들의 가중 선형 조합으로 이루어지며 식(24)와 같다.

$$\hat{y} = \sum_{i=1}^C \sum_{j=1}^r u_j^i \cdot w_j^i \quad (24)$$

여기서,  $u_j^i$ 는  $i$ 번째 context의  $j$ 번째 클러스터의 멤버십 함수의 소속 정도를 의미하며,  $w_j^i$ 는  $i$ 번째 context의  $j$ 번째 클러스터와 출력 노드간의 연결하중을 의미한다.

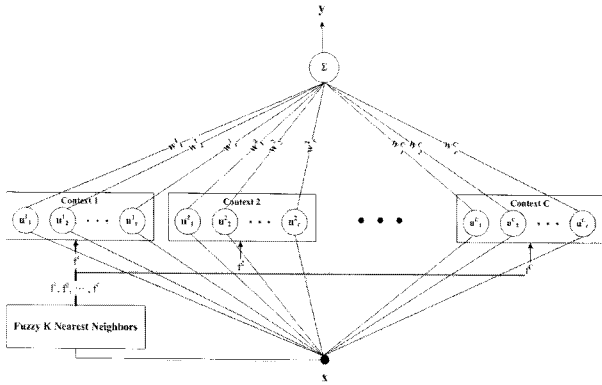


그림 3 제안된 RBF 뉴럴 네트워크의 구조  
 Fig. 3 The architecture of the proposed RBF neural networks

테스트 집합의 부가적 정보 입자를 추출하고 추출된 정보 입자를 기반으로 c-FCM을 적용하여 RBF를 정의한 RBF 뉴럴 네트워크의 설계 과정은 아래와 같다.

- Step 1. 출력공간상에서 정의된 부가적 정보 입자( $f'_k$ ) 추출
- Step 2. FKNN기법을 이용하여 테스트 데이터 집합의 부가적 정보입자들의 적합도 추정
  - Sub-Step 2.1 j번째 테스트데이터를 query instance로 정의
  - Sub-Step 2.2 k개의 nearest neighbor로 추출
  - Sub-Step 2.3 정의된 k nearest neighbors와 (25)를 이용하여 query instance의 1번째 부가적 정보입자의 적합도 추정

$$\hat{f}^j(q) = \sum_{i=1}^K S(x_{L(i)}, q) \cdot f'_i = \sum_{i=1}^K \left( \frac{f'_i}{\sum_{j=1}^K \left( \frac{\|x_{L(i)} - q\|}{\|x_{L(j)} - q\|} \right)^{\frac{2}{p-1}}} \right) \quad (25)$$

- Sub-Step 2.4 sub-step 2.1에서 sub-step 2.3까지의 과정을 모든 테스트 데이터 집합의 데이터에 관하여 수행한다.
- Step 3. 학습 데이터와 테스트 데이터 집합의 partition matrix를 구한다.
- Step 4. Least Square Estimation을 이용하여 RBF 뉴럴 네트워크의 후반부 계수인 연결 하중 추정

### 5. 실험 연구 및 결과 고찰

Radial Basis Function Neural Networks의 성능을 개선하기 위한 방법으로 conditional fuzzy C-means 클러스터링을 이용한 RBFNN의 receptive field의 위치를 설정하는 방법이 Pedrycz에 의해 제안되었다. Pedrycz에 의해 제안된 방법은 conditional fuzzy c-means 클러스터링의 장점인 출력공간상의 정보 입자들의 정보의 활용이라는 특징을 최대한 이용하지 못하고 있다. 이러한 단점을 극복하기 위하여 본 논문에서는 fuzzy K-Nearest Neighbors 기법을 이용하여 테스트 집합에서도 출력 공간상에 정의된 부가적 정보 입자들의 정보를 활용할 수 있도록 개선하였다. 제안된 퍼지 KNN과 CFM을 이용하여 설계된 RBFNN의 성능 평가를 위하여 본 논문에서는 두가지 종류의 machine learning 데이터 집합 (1: Automobile MPG (Miles Per Gallon) Data, 2: Boston Housing Data)에 제안된 알고리즘을 적용하여 성능을 평가하였다. 성능 평가를 위한 평가 지수는 Root Mean Square Error(RMSE)를 사용하였다.

$$PI = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (26)$$

제안된 알고리즘의 성능평가를 위한 실험조건은 표 1과 같다.

표 1 제안된 모델의 실험 파라미터  
 Table 1 Selected numeric values of the parameters

Parameter	Value
Polynomial Order (O)	1 (linear)
Number of Nearest Neighbors (K)	2 ~ 7
Fuzzification Coefficient (p)	1.2 ~ 4.0
Number of Contexts (C)	2 ~ 6
Number of Rules per each context (r)	2 ~ 6

#### 5.1 Automobile MPG (Miles Per Gallon) Data

Automobile MPG 392개의 데이터 쌍들 중에서 60%를 학습 데이터 집합으로 설정하고 나머지 40%를 테스트 데이터 집합으로 임의적으로 분할한다. 이와 같은 임의적 분할을 10번 수행하고, 임의로 분할된 학습 데이터와 테스트 데이터를 이용하여 제안된 모델 설계 방법의 성능을 평가한다.

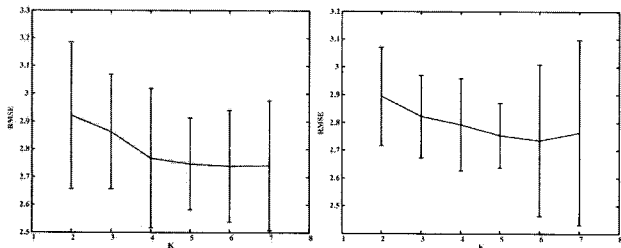
표 2는 Nearest Neighbor의 수(K)의 개수에 따라 가장 좋은 성능을 보이는 모델의 파라미터와 성능 지수를 나타낸다. 표 2에 명시된 K는 nearest neighbor의 수, p는 fuzzification coefficient, C는 출력공간상에서 정의된 context의 수, R은 각 context와 관련된 입력공간상에서의 퍼지 규칙의 수를 의미한다. 표 2는 K-Means 클러스터링 방법을 이용하여 얻는 모델의 성능지수를 보인다.

표 2 제안된 모델의 성능 지수  
 Table 2 The performance index of the proposed model

K	C	r	p	Training Data		Test Data	
				mean	STD	mean	STD
2	5	5	2.4	1.2151e-14	2.1775e-15	2.8938	0.1761
3	3	2	2	1.1132e-14	4.7696e-15	2.8218	0.1486
4	3	5	3.4	1.3217e-14	7.1964e-15	2.7932	0.1656
5	3	4	2.4	1.1224e-14	3.3730e-15	2.7532	0.1159
6	4	4	2.2	1.0881e-14	4.5571e-15	2.7355	0.2726
7	2	2	1.6	8.6753e-15	3.4945e-15	2.7638	0.3327

표 2에서 보인 모델의 성능 지수들 중 가장 선호되는 모델은 K=6, C=4, r=4, p=2.2인 RBFNN이며 출력공간상에 정의된 context는 K-Means 클러스터링에 의해 정의된다. 선호된 모델의 테스트 데이터 집합에 대한 모델링 성능(EPI)은 2.7355±0.2726이다. 그림 4는 Context를 추출하기 위한 방법으로 Evenly partition (a) 과 K-Means 클러스터링 (b)방법을 사용하여 설계한 모델의 성능지수를 K값에 따라 도시한 것이다.

그림 4에서 보인 결과를 분석하면, K값이 커짐에 따라 모델의 성능이 대체적으로 개선됨을 알 수 있다. 그림 7은 fuzzification coefficient(p) 값의 변화에 따른 모델의 테스트 데이터에 대한 성능지수를 보인다. 그림 5에 도시한 모델은 context의 수가 4이며, 근접한 이웃의 수(K)가 6이고, context 추출 방법이 K-Means 클러스터링을 사용한 경우에 p값의 변화에 따른 모델의 성능지수의 변화를 보인 것이다.



(a) Evenly Partition (b) K-Means 클러스터링

그림 4 K값에 따른 모델의 성능지수

Fig. 4 Performance Index versus the number of the nearest neighbors

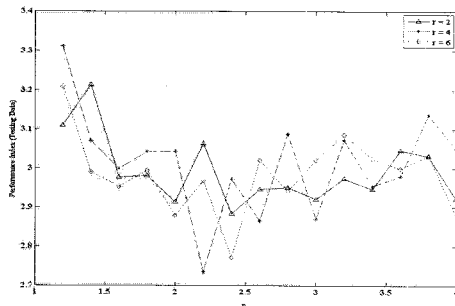


그림 5 퍼지화 상수 p값의 변화에 따른 모델의 성능지수

Fig. 5 The performance index of the proposed model versus the variation of fuzzification coefficient (p)

표 3은 제안된 모델과 기존의 다른 모델과의 모델링 성능 비교를 보인다.

표 3 모델의 성능 비교

Table 3 Results of comparative analysis

Model	parameters	PI(RMSE)	EPI(RMSE)
RBFNN [17]	C=36	3.24±0.24	3.62±0.31
RBFNN with context-free clustering [17]	C=36	3.21±0.21	3.51±0.27
RBFNN*	C=9	7.037±0.153	7.615±0.321
RBFNN with FCM	C=27	3.712±0.130	3.522±0.314
Single PNN	N/A	2.563±0.146	2.948±0.348
MARS	N/A	2.726±0.103	3.0523±0.211
Proposed Model	Evenly Partition	K=6, C=5, r=4, p=2.2 1.06e-14 ±3.047e-15	2.739±0.201
	K-Means Clustering	K=6, C=4, r=4, p=2.2 1.088e-14 ±4.557e-15	2.736±0.273

표 3에서, PI는 학습 데이터에 대한 모델링 성능을 의미하며, EPI는 테스트 데이터에 대한 모델링 성능을 의미한다. RBFNN\*는 MATLAB에서 제공하는 toolbox를 이용하여 얻은 결과이다.

5.2 Boston Housing Data

Boston Housing Data 집합은 보스턴 지역의 실제 부동산에 대한 정보를 취급하는 데이터들의 집합이다. 506개의 데이터 집합을 60%와 40%의 데이터 집합으로 랜덤하게 분할한다. 10번 반복하여 랜덤 분할하여 제안된 모델의 성능을 평가한다. 위와 같은 방법으로 평가된 제안된 모델의 성능 지수는 표 4에 나열하였다. 표 4에서 보인 모델의 성능 지수들 중 가장 선호되는 모델은 K=6, C=4, r=5, p=1.4인

RBFNN이며 출력공간상에 정의된 context는 K-Means 클러스터링에 의해 정의된다. 선호된 모델의 테스트 데이터 집합에 대한 모델링 성능(EPI)은 3.7257±0.2718이다.

표 4 제안된 모델의 성능 지수

Table 4 The performance index of the proposed model

K	C	r	p	Training Data		Test Data	
				mean	STD	mean	STD
2	4	2	3.6	1.528e-14	2.663e-15	3.8379	0.5173
3	6	2	1.6	1.558e-14	6.800e-15	3.8358	0.491
4	6	2	2.4	1.286e-14	3.921e-15	3.8086	0.6158
5	4	3	1.4	1.547e-14	8.622e-15	3.8452	0.3793
6	4	5	1.4	1.158e-14	3.285e-15	3.7257	0.2718
7	3	2	1.8	1.441e-14	6.292e-15	3.8109	0.4742

표 5는 제안된 모델과 기존의 다른 모델과의 모델링 성능 비교를 보인다. 표 5에서, PI는 학습 데이터에 대한 모델링 성능을 의미하며, EPI는 테스트 데이터에 대한 모델링 성능을 의미한다. RBFNN\*는 MATLAB에서 제공하는 toolbox를 이용하여 얻은 결과이다.

표 5 모델의 성능 비교

Table 5 Results of comparative analysis

Model	parameters	PI(RMSE)	EPI(RMSE)
RBFNN [17]	C=25	6.36±0.24	6.94±0.31
RBFNN with context-free clustering [17]	C=25	5.52±0.25	6.91±0.45
RBFNN*	C=18	6.751±0.279	9.123±0.532
RBFNN with FCM	C=22	6.771±0.383	6.581±0.442
Single PNN	N/A	3.722±0.359	5.446±3.461
MARS	N/A	3.255±0.225	4.188±0.330
Proposed Model	Evenly Partition	K=4, C=2, r=4, p=1.8 1.212e-14 ±3.709e-15	3.759±0.284
	K-Means Clustering	K=6, C=4, r=5, p=1.4 1.158e-14 ±3.285e-15	3.726±0.272

6. 결론

본 논문에서는 RBFNN의 receptive field들의 위치 설정을 위한 알고리즘으로 많이 연구되어지고 있는 conditional fuzzy c-means 클러스터링의 단점인 출력 공간상에서의 부가적 정보입자들의 추정을 위하여 fuzzy K-nearest neighbors 기법을 적용하였다. Conditional fuzzy c-means 클러스터링 기법은 출력 공간상에서 출력변수들의 분포를 분석하여 입력 공간에 receptive field의 위치를 설정하여, RBFNN의 모델링 성능을 개선시킨다. 그러나 출력변수의 값이 알려져 있지 않은 테스트 데이터의 경우에는 출력공간상에서의 부가적 정보입자를 얻을 수 없기 때문에 conditional fuzzy c-means 클러스터링을 적용할 수 없게 된다. 이를 극복하기 위하여 fuzzy k-nearest neighbor 기법을 적용하여 출력변수가 알려져 있지 않은 테스트 데이터 집합에도 conditional fuzzy c-means 클러스터링을 적용할 수 있도록 한다.

제안된 알고리즘에 의해 설계된 RBFNN은 기존 모델에 비해 학습 데이터의 모델링 성능 지수가 매우 우수한 것을 알 수 있었다. 더불어 제안된 모델의 일반화 능력 (테스트 데이터에 대한 모델링 성능) 또한 기존의 다른 모델에 비해 우

수한 것을 보인다. 기존 모델에 비해 제안된 모델은 일반화 성능(EPI)면에서의 성능개선 보다는 근사화 성능(PI) 측면에서의 성능개선이 두드러진 특징으로 나타난다. 이러한 현상은 제안된 모델의 RBF 정의에 출력 공간상에서 정의된 context 정보가 반영되어지기 때문이다 (식 (12) 참조). 그러나 테스트 데이터의 경우에는 정확한 context 정보를 확보할 수 없기 때문에 FKNN을 이용하여 추정된 context의 정보를 이용한다. 추정된 context의 정보가 정확하면 정확 할수록 제안된 모델의 일반화 성능(EPI)도 더욱 개선될 것으로 분석된다.

**감사의 글**

이 논문은 2006년도 정부(교육인적자원부)의 재원으로 한국 학술진흥 재단의 지원을 받아 수행된 연구임 (KRF-2006-353-D00021) 그리고 이 논문은 2008년도 정부(교육과학기술부)의 재원으로 한국학술진흥 재단의 지원을 받아 수행된 연구임 (KRF-2008-314-D00376).

**참 고 문 헌**

[1] M.J.D. Powell, Radial basis functions for multivariable interpolation: a review, in: J.C. Mason, M.G. Cox (Eds.), Algorithms for Approximation, Oxford University Press, Oxford, 1987, pp. 143-167.  
 [2] W. Pedrycz, H.S. Park, S.K. Oh, A granular-oriented development of functional radial basis function neural networks, Neurocomputing 72 (2008) 420-435  
 [3] Wing W.Y. Ng, Andres Dorado, Daniel S. Yeung, Witold Pedrycz, Ebroul Izquierdo, Image classification with the use of radial basis function neural networks and the minimization of the localized generalization error, Pattern Recognition 40 (2007) 19-32  
 [4] Carlos Renjifo, David Barsic, Craig Carmen, Kevin Norman, G.Scott Peacock, Improving radial basis function kernel classification through incremental learning and automatic parameter selection, Neurocomputing 72 (2008) 3-14  
 [5] S. Albrecht, et al., Generalized radial basis function networks for classification and novelty detection: self-organization of optimal Bayesian decision, Neural Networks 13 (2000) 1075-1093.  
 [6] G. Bugmann, Normalized Gaussian radial basis function networks, Neurocomputing 20 (1998) 97 - 110.  
 [7] C. G. Looney, Radial basis functional link nets and fuzzy reasoning, Neurocomputing 48 (2002) 489 - 509.  
 [8] L. Marcelino, S. Ignacio, P. Carlos, A new EM-based training algorithm for RBF networks, Neural Networks 16 (1) (2003) 69-77.  
 [9] A. Staiano, R. Tagliaferri, W. Pedrycz, Improving RBF networks performance in regression tasks by means of a supervised fuzzy clustering, Neurocomputing 69 (13-15) (2006) 1570-1581.  
 [10] L. Ma, A. Wahab, G. S. Ng, S. Erdogan, An

experimental study of the extended NRBF regression model and its enhancement for classification problem, Neurocomputing 72 (2008) 458-470.  
 [11] A. Staiano, R. Tagliaferri, W. Pedrycz, Improving RBF networks performance in regression tasks by means of a supervised fuzzy clustering, Neurocomputing 69 (13 - 15) (2006) 1570-1581  
 [12] S. Mitra, J. Basak, FRBF: a fuzzy radial basis function network, Neural Comput. Appl. 10 (2001) 244-252.  
 [13] W. Pedrycz, Conditional fuzzy clustering in the design of radial basis function neural networks, IEEE Trans. Neural Netw. 9 (4) (1998) 601 - 612.  
 [14] N. B. Karayiannis, Reformulated radial basis neural networks trained by gradient descent, IEEE Trans. Neural Netw. 10 (3) (1999) 657-671.  
 [15] S. M. Loone, G. Irwin, Improving neural network training solutions using regularization, Neurocomputing 37 (2001) 71-90.  
 [16] L. Xu, RBF nets, mixture experts, and Bayesian Ying-Yang learning, Neurocomputing 19 (1998) 223-257.  
 [17] W. Pedrycz and K. C. Kwak, "Linguistic Models as a Framework of User-Centric System Modeling", IEEE Trans. on Systems, Man and Cybernetics-A, Vol. 36, No. 4, pp. 727-745, 2006

**저 자 소 개**



**노 석 범 (盧石範)**

1994년 원광대 제어계측공학과 졸업.  
 1996년 동 대학 컴퓨터 공학과 졸업(석사). 2006년 동 대학 제어계측공학과 졸업(박사). 관심 분야 : 퍼지 모델링, 컴퓨터 지능, statistical learning  
 Tel : (063) 850-6344  
 E-mail : nado@wonkwang.ac.kr



**오 성 권 (吳聖權)**

1981년 연세대 전기공학과 졸업. 1983-89년 금성산전연구소(선임연구원). 1983년 연세대 대학원 전기공학과 졸업(공학박사). 1996-1997년 캐나다 Manitoba대학 전기 및 컴퓨터공학과 Post-Doc. 1983-2003년 원광대학교 전기전자 및 정보공학부 교수. 2005년-현재 수원대학교 정교수. 관심 분야 : 시스템 자동화, 퍼지 이론 및 뉴로 퍼지 네트워크 응용, 계산지능 및 소프트 컴퓨팅  
 Tel : (031) 229-8162  
 E-mail : ohsk@suwon.ac.kr