

Context-based 클러스터링에 의한 Granular-based RBF NN의 설계

논문

58-6-26

The Design of Granular-based Radial Basis Function Neural Network by Context-based Clustering

박 호 성^{*} · 오 성 권^{*}
(Ho-Sung Park · Sung-Kwun Oh)

Abstract - In this paper, we develop a design methodology of Granular-based Radial Basis Function Neural Networks (GRBFNN) by context-based clustering. In contrast with the plethora of existing approaches, here we promote a development strategy in which a topology of the network is predominantly based upon a collection of information granules formed on a basis of available experimental data. The output space is granulated making use of the K-Means clustering while the input space is clustered with the aid of a so-called context-based fuzzy clustering. The number of information granules produced for each context is adjusted so that we satisfy a certain reconstructability criterion that helps us minimize an error between the original data and the ones resulting from their reconstruction involving prototypes of the clusters and the corresponding membership values. In contrast to "standard" Radial Basis Function neural networks, the output neuron of the network exhibits a certain functional nature as its connections are realized as local linear whose location is determined by the values of the context and the prototypes in the input space. The other parameters of these local functions are subject to further parametric optimization. Numeric examples involve some low dimensional synthetic data and selected data coming from the Machine Learning repository.

Key Words : Granular-based RBF NNs, Receptive fields, Fuzzy clustering, Context-based fuzzy clustering, Information granules

1. 서 론

뉴로 컴퓨팅분야에서 Radial Basis Function (RBF) 뉴럴 네트워크가 도래한 후, 많은 관심과 더불어 그와 같은 형태의 뉴럴 네트워크의 일반적인 형태의 유용한 확장을 이룰 수 있었다. 즉, RBF 뉴럴 네트워크에 대한 다수의 비교적인 연구와 다양한 적용이 이루어지고 있다[1-3]. 방사형 기저 함수들로 이루어져 있는 RBF 뉴럴 네트워크의 일반적인 구조는 출력 뉴런에 의해서 선형적인 방법으로 결합되는 구조를 가지고 있으며, 다음과 같이 몇 개의 주요한 개발 방향과 설계 전략을 가지고 있다.

(a) 전형적인 형태의 기저함수로서 가우시안 함수를 이용한 RBF 뉴럴 네트워크가 다양한 분야에서 사용되어져 왔으며 여러 가지 서로 다른 방법들로 개발되어져 왔다[4-7]. 또한 어떠한 주어진 실험적인 데이터를 기초로 해서 은닉층에서의 기저함수를 K-Means와 Fuzzy C-Means 클러스터링 방법[8]과 같이 표현되어지는 다양한 클러스터링 기법을 사용을 통해 대안적인 방법으로 처리하는 방법도 널리 사용되어지고 있다[9-12].

(b) RBF 뉴럴 네트워크 안에서 수행된 파라미터의 최적화를 위한 다양한 학습 방법들이 제안되어지고 있다. 즉, 은닉층에서의 파라미터뿐만 아니라 네트워크의 출력층에서 위치한 선형 뉴런들의 파라미터들의 추정하기 위하여 다양한 학습 방법들이 사용된다. 다시 말해서 경사 하강법 뿐만 아니라 진화론적 최적화를 포함한 전역 극소 기법들이 시도되었다[13-14].

본 연구에서는 정보 입자화의 세심한 처리를 통한 실험적 데이터의 효과적인 사용을 기반으로 처리하는 탁월한 데이터 유도 구조로서의 Granular-based RBF 뉴럴 네트워크 (GRBFNN)를 제안한다. 본 논문에서의 궁극적인 목적은 (a) 퍼지 클러스터링의(context-driven Fuzzy C-Means와 관련된) 절차에 따라 출력변수에 의해 구현된 퍼지 입자화의 사용, (b) 입력 공간의 특성과 신중하게 형성된 정보 입자화를 통한 입력공간의 균형 잡힌 표현을 나타내도록 하는 각각의 context에 대한 클러스터의 조절 가능한 수의 고려, (c) 네트워크의 출력층에서 표현된 뉴런의 함수적 연결의 표현, 그리고 (d) 다양한 퍼지 계수(m)에 따른 모델의 효율적인 구축과 같은 GRBFNN의 포괄적인 설계 방법론을 제안하는 것이다.

본 논문에서 제안한 GRBFNN은 퍼지 집합과 뉴로-컴퓨팅 사이의 시너지 효과를 나타내는 뉴로-퍼지 구조의 한 분야에 포함되어 있으며, 이 시너지효과는 네트워크의 형태가 개발될 때 구조적 측면에서 드러나게 된다. 또한 구조적 개발의 다양성을 입증하고 제안된 구조의 장점과 한계를 정량

^{*} 교신저자, 정회원 : 수원대 산업기술연구소 학술연구교수 · 공박

E-mail : parkhs@suwon.ac.kr

^{*} 정 회 원 : 수원대 공대 전기공학과 교수 · 공박

접수일자 : 2009년 3월 11일

최종완료 : 2009년 5월 15일

화하기 위하여 저차원의 합성 데이터와 Machine learning 데이터 집합들을 실험적 데이터로 이용하였다.

2. Context-based 클러스터링의 사용을 통한 데이터 안에서의 입자화 구조의 표현

2.1 K-Means 클러스터링을 이용한 Context fuzzy set 구성

K-Means 클러스터링은 영상 데이터의 압축[15]과 시스템 모델링을 위한 데이터 전처리[16]와 같은 다양한 영역에서 적용되어져 왔으며, 데이터의 구성과 분류뿐만 아니라 데이터의 압축과 모델의 동정에도 널리 사용되어져 왔다.

본 논문에서는 이러한 K-Means 클러스터링 알고리즘을 이용하여 출력변수인 $\{target_k\}$ 을 P-2개로 분류하고 다음 식 (1)과 같이 어떠한 목적함수 값 Q가 최소가 되도록 각각의 클러스터의 중심값을 찾는다.

$$Q = \sum_{i=2k=1}^{P-1} \sum_{i=1}^N (target_k - y_i)^2 \quad (1)$$

또한 출력변수의 최소 및 최대값을 첫 번째와 마지막 클러스터의 중심값으로 설정한다. 따라서 K-Means 클러스터링 알고리즘을 이용하여 전체적으로 P개의 중심이 생성된다. 이와 같이 얻어진 중심값들은 각각 y_1, y_2, \dots, y_P 로 표기하고 그림 1과 같이 삼각형 멤버십 함수의 중심값으로 이용한다. 삼각형 멤버십 함수의 형태는 두 개의 연속적인 퍼지 집합의 겹침도를 1/2이 되도록 멤버십 함수를 구성한다.

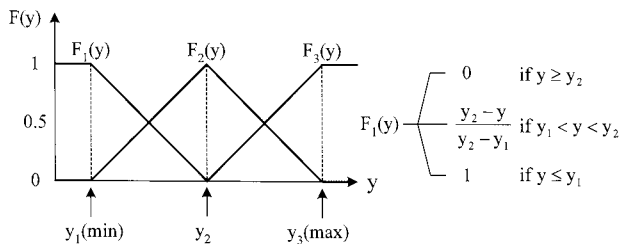


그림 1 삼각형 멤버십 함수와 파라미터 (P=3)
Fig. 1 Triangular membership functions and their parameters (P=3)

이와 같은 방법을 통해 얻어진 멤버십 함수의 값을 F_1, F_2, \dots, F_P 로 각각 표기하며, 특별히 j-번째 context fuzzy set에서의 데이터 $\{target_k\}$ 의 멤버십 정도를 f_{jk} 로 표기한다.

2.2 Context-based Fuzzy C-Means 클러스터링

Context-based FCM 클러스터링은 출력공간에서 몇 개의 미리 정의된 context라 불리는 퍼지 집합의 분류들을 기준으로 입력 데이터에서의 공간을 정보 입자화의 특성에 맞게 분류하는 알고리즘이다. [10, 17, 18]. 이와 같이 context를 고려함으로써, 입력공간에서의 클러스터는 몇 개의 미리 정의된 context의 퍼지 집합에 의해서 분류되어져 간다. 이는

입력공간과 출력 공간 영역에서 사이의 관계를 드러나게 함으로써, 기존의 입력만을 고려한 클러스터링 방법에서 벗어나 입력과 출력을 동시에 고려한 특징을 가지고 있다.

Context-based FCM 클러스터링 방법으로 입력 데이터 $\{x_k\}$ 를 'c'개의 클러스터로 분할한다. $U(F_j)$ 라 표시되는 j-번째 context에 의해서 유도되는 소속 행렬을 구하기 위한 context-based FCM 클러스터링 알고리즘은 다음과 같다.

$$U(F_j) = \left\{ u_{ik} \in [0, 1], \sum_{i=1}^c u_{ik} = f_{jk} \forall k, 0 < \sum_{k=1}^N u_{ik} < N \forall i \right\} \quad (2)$$

여기서, c는 입력 데이터에서의 클러스터의 수, f_{jk} 는 삼각형 멤버십 함수에 의해서 얻어진 j-번째 context에 속해 있는 k-번째 데이터들의 멤버십 값을 의미한다.

Context-based FCM 클러스터링의 목적함수는 다음과 같다.

$$V = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|x_k - z_i\|^2 \quad (3)$$

여기서, m>1은 퍼지 계수를 나타내고, z_i 는 입력변수들의 중심값을 의미한다.

목적함수 V는 식 (2)에서 표현된 조건에 만족하도록 최소화되며, 이는 각각의 context들에 의해 분류된 'P'개의 클러스터링에 적용된다. 간단히 다음과 같이 쓸 수 있다.

$$\text{Min } V \text{ subject to } U(F_j), \quad j = 1, 2, \dots, P \quad (4)$$

Context-based FCM에 의해서 실행되는 목적함수 V의 최소화는 partition matrix와 중심값들이 반복적으로 갱신되는 것으로 실현된다.

Partition matrix는 다음과 같이 실행된다.

$$u_{ik} = \frac{f_{jk}}{\sum_{i=1}^c \left(\frac{\|x_k - z_i\|}{\|x_k - z_l\|} \right)^{\frac{2}{m-1}}}, \quad i = 1, 2, \dots, c, k = 1, 2, \dots, N \quad (5)$$

식 (5)에서 볼 수 있듯이, 기존의 FCM 클러스터링과 다른 특징은 식 (5)의 분자식이 '1'이 아닌 식 (1)에서 얻은 출력공간에서의 멤버십 값이라는 것이다. 이는 입력공간에서의 정보 입자화를 해당되는 context를 기반으로 입력공간을 분류하는 것으로서, 입력과 출력과의 관계를 보여준다.

중심값 (z_1, z_2, \dots, z_c)은 다음과 같은 형태로 이루어진다.

$$z_i = \frac{\sum_{k=1}^N u_{ik}^m x_k}{\sum_{k=1}^N u_{ik}^m} \quad (6)$$

여기서 $i=1, 2, \dots, c$.

모든 context들에 대해서 context-based FCM 클러스터링을 수행한다면, 다음과 같은 prototype들이 생성된다.

$$v_1 v_2 \dots v_c v_{c+1} \dots v_{cP} \quad (7)$$

여기서 식 (7)에서 첫 번째 "c" 요소는 첫 번째 context인 F1에 의해서 생성된 입력변수들의 중심값들을 의미하고 다음의 "c" 요소는 두 번째 context에 관련된 입력변수들의 중심값을 나타낸다. 따라서 최종적으로는 입력 공간에서 "cP" 개의 클러스터가 형성된다.

위에서 나타낸 것과 같이 열거된 중심값들의 인덱스를 L(j)={i1, i2, ..., ic}와 같이 표기하면 다음과 같이 L(1)={1, 2, ..., c}, L(2) = {c+1, c+2, ..., 2c} 등과 같이 결론적으로 연속적인 정수의 수들의 집합의 형태로 표현할 수 있다.

지금까지 개발된 클러스터의 구조는 다음 그림 2에서 간략하게 보여준 바와 같이 조직적으로 표현할 수 있다.

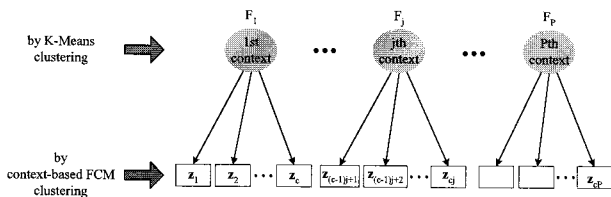


그림 2 출력과 입력 공간에서 형성된 입자화 구조의 구조
Fig. 2 A structure of granular constructs formed in the output and input spaces

3. Reconstructability criterion에 의한 입자화 구조의 개선

기존의 context-based FCM 클러스터링 방법을 이용한 모델들[10, 17, 18]은 각각의 context에 똑같은 수의 클러스터를 적용하였다. 이는 데이터의 서로 다른 구조 및 복잡성에도 불구하고 서로 다른 context에 똑같은 수의 클러스터를 적용하는 단점을 가지고 있다. 본 논문에서는 이와 같은 구조의 단점을 개선하기 위해서 각각의 context에서 필요할 때까지 각각의 context에 대해서 클러스터의 수를 증가하는 것을 고려한다. 이는 본질적으로 각각의 context에 대한 클러스터의 수는 다양할 수 있을 것이라는 가정으로 기초로 하며 본 논문에서 다음과 같이 제안한 reconstructability criterion 알고리즘을 기반으로 각각의 context에 대한 클러스터의 수를 조절한다.

먼저 각각의 입력변수 x_k 에 대해서 다음과 같이 클러스터의 인덱스 (i)를 결정한다.

$$i = \arg \max u_{ik} \quad (8)$$

여기서 u_{ik} 는 다음과 같이 계산된다.

$$u_{ik} = \frac{1}{\sum_{l=1}^c \left(\frac{\|x_k - v_l\|}{\|x_k - v_i\|} \right)^{2/(m-1)}} \quad (9)$$

또한 식 (10)과 같이 표현하는 것과 같은 관계를 가지는 데이터의 부분 집합을 Ω_j 로 표기한다.

$$\Omega_j = \{x_k \mid i = \arg \max u_{ik} \text{ and } i \in L(j)\} \quad (10)$$

모든 $x_k \in \Omega_j$ 는 j-번째 context에 의해서 포함된 중심값들을 기반으로 하여 일정한 정확성을 가지고 "재구성"될 수 있다. 즉, 재구성된 데이터 \hat{x}_k 는 다음과 같이 계산된다.

$$\hat{x}_k = \frac{\sum_{i \in L(j)} g_{ik} v_i}{\sum_{i \in L(j)} g_{ik}} \quad (11)$$

여기서, 소속 정도 g_{ik} 는 j-번째 context에 의해 포함된 중심값들을 기초로 결정된다. 즉,

$$g_{ik} = \frac{1}{\sum_{i \in L(j)} \left(\frac{\|x_k - v_i\|}{\|x_k - v_i\|} \right)^{2/(m-1)}} \quad (12)$$

여기서, $i \in L(j)$ 이고 $x_k \in \Omega_j$.

식 (13)에서 보이듯이, 원 데이터 x_k 와 재구성된 데이터 \hat{x}_k 와의 weighted Euclidean distance 계산을 통해서 재구성의 특징인 평가기준을 측정한다.

$$V(j) = \sum_{k \in \Omega_j} \|x_k - \hat{x}_k\|^2 \quad (13)$$

식 (13)에서 제안된 평가기준의 계산은 모든 context에 대해서 실행되면 이는 결과적으로 V(1), V(2), ..., V(P)와 같은 결과값으로 생성된다. 여기서 우리는 j0-번째 context에서 발생한 가장 작은 값을 가진 V~값을 선택하고 이를 기준으로 간주한다. 그 후 j0-번째 context를 제외한 나머지 context에 대해서 클러스터의 수를 증가하여 V(j)의 값이 V~값보다 작거나 또는 같을 때까지 클러스터의 수를 증가시키면서 context-based FCM 클러스터링 알고리즘을 계산한다. 만일 필요하다면 클러스터의 수를 계속 증가시키고 context-based FCM 클러스터링을 실행한다. 단, 모델의 복잡성을 고려하여 본 논문에서는 클러스터의 수 증가에 대한 강제 제한 조건으로 '2c'를 한정하였는데 이는 클러스터의 수 증가가 기준인 V~의 값을 가지는 클러스터의 수보다 2배가 초과하지 않도록 하기 위함이다.

4. 제안된 Granular-based RBF Neural Network

앞장에서 언급되어진 알고리즘 및 구조를 기반으로 본 논문에서 제안한 GRBFNN에 대한 구조는 그림 3에서 보여진바와 같다. 그림 3에서도 볼 수 있듯이 본 논문에서 제안한 모델은 각각의 context에 적당한 수의 클러스터를 조절함으로써 주어진 데이터의 정보 입자화를 모델에 효과적으로 반영하도록 하였다.

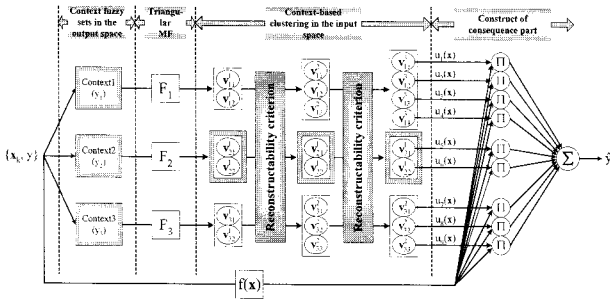


그림 3 제안된 Granular-based RBF 뉴럴 네트워크의 전체 구조

Fig. 3 The overall structure of the Granular-based RBF Neural Network

은닉층으로부터의 연결은 함수적인 특성을 지니고 있으며, 지역 모델은 식 (14)에서 볼 수 있듯이 입력과 출력에서 형성된 정보 입자화 주위에 선형 형태로 이루어져 있다.

$$h - y_l = \mathbf{a}_l^T (\mathbf{x} - \mathbf{v}_l) \quad (14)$$

여기서 \mathbf{v}_l 는 $l \in L(I)$ 인 l -번째 context에 위치한 입력공간의 중심값이고, y_l 은 l -번째 context의 퍼지 집합의 중심값이다. "h"는 출력변수를 의미한다.

제안된 모델의 전체 출력은 식 (15)과 같이 모두 (C)개의 지역 모델의 가중된 평균에 의해서 계산된다.

$$y(\mathbf{x}) = \sum_{i \in L(1)} u_i(\mathbf{x}) [y_1 + \mathbf{a}_1^T (\mathbf{x} - \mathbf{v}_1)] + \sum_{i \in L(2)} u_i(\mathbf{x}) [y_2 + \mathbf{a}_2^T (\mathbf{x} - \mathbf{v}_2)] + \dots + \sum_{i \in L(P)} u_i(\mathbf{x}) [y_p + \mathbf{a}_i^T (\mathbf{x} - \mathbf{v}_i)] \quad (15)$$

여기서 소속 정도 $u_i(\mathbf{x})$ 는 지금까지 형성된 클러스터들의 모든 중심값들을 고려하여 식 (16)과 같이 계산된다.

$$u_i(\mathbf{x}) = \frac{1}{\sum_{l=1}^C \left(\frac{\|\mathbf{x}_k - \mathbf{v}_l\|}{\|\mathbf{x}_k - \mathbf{v}_i\|} \right)^{2/(m-1)}} \quad (16)$$

여기서, $i=1, 2, \dots, C$.

지역 모델들의 파라미터 \mathbf{a}_i 의 최적화는 식 (17)이 최소화될 수 있도록 standard Mean Square Error 이용하였다.

$$E = \sum_{k=1}^N (y(\mathbf{x}_k) - \text{target}_k)^2 \quad (17)$$

5. 시뮬레이션 및 결과고찰

몇 개의 실험적 데이터를 이용하여 제안된 설계 방법론의 해석에 대한 탁월한 능력의 우수성을 입증하고자 한다. 첫 번째 데이터 집합은 3차원의 합성 데이터 집합이고, 두 번째 데이터 집합은 Machine Learning repository로부터 구한 몇 개의 선택된 데이터 집합을 이용한다. 모든 실험들은 10-fold cross validation 모드를 실행하였다.

5.1 Synthetic Three-Dimensional Data

첫 번째 데이터 집합으로써 식 (18)에서 주어진 것과 같이 3입력 추력의 비선형 함수식을 고려한다.

$$y = (1 + x_1^{-2} + x_2^{-1.5} + x_3^{-1.0})^2, \quad 1 \leq x_1, x_2, x_3 \leq 5 \quad (18)$$

식 (18)을 이용하여 총 200개의 입출력 쌍을 생성한다. 전체 데이터 집합으로부터 랜덤하게 학습 데이터 집합(전체 데이터 집합중의 60%)을 선택하고 나머지 데이터 집합들은 테스트 데이터 집합으로 사용한다. 또한 다양한 정보 입자화의 효과를 평가하기 위하여 context의 수를 2~6개 사이로 정하여 다양하게 실험 및 측정하였다.

표 1은 각각의 context에서 사용된 퍼지 계수(m)에 따른 제안된 모델의 각각의 context의 성능을 요약한 것이다.

표 1 3차원 합성 데이터 대한 성능지수

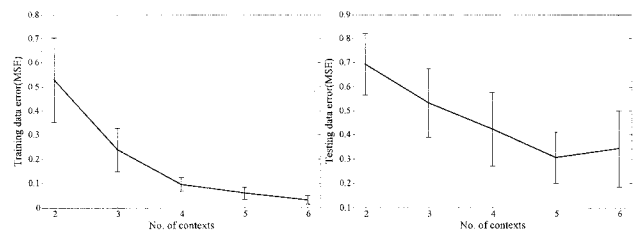
Table 1 Values of the performance index for the synthetic-three dimensional data : PI-performance index for the training data, EPI-performance index for the testing data

(a) Values of the PI

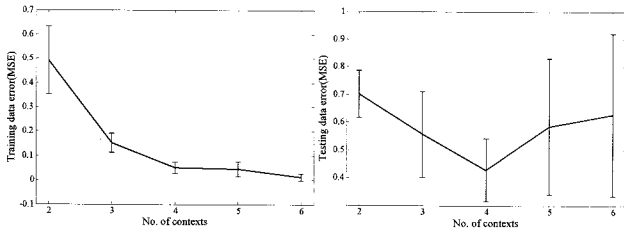
m	Number of contexts(P)				
	2	3	4	5	6
1.1	0.367±0.123	0.260±0.072	0.130±0.083	0.147±0.189	0.064±0.036
1.5	1.198±0.475	0.376±0.144	0.162±0.030	0.067±0.040	0.082±0.067
2.0	0.491±0.140	0.152±0.038	0.050±0.023	0.059±0.025	0.031±0.015
2.5	0.579±0.103	0.347±0.040	0.211±0.076	0.144±0.058	0.111±0.104
3.0	0.489±0.075	0.378±0.061	0.338±0.045	0.270±0.076	0.225±0.067
3.5	0.442±0.057	0.361±0.059	0.347±0.041	0.274±0.063	0.255±0.069
4.0	0.452±0.077	0.382±0.049	0.355±0.037	0.283±0.062	0.230±0.044
4.5	0.418±0.079	0.373±0.048	0.353±0.061	0.277±0.072	0.250±0.030
5.0	0.446±0.054	0.386±0.073	0.359±0.034	0.307±0.037	0.272±0.049

(b) Values of the EPI

m	Number of contexts(P)				
	2	3	4	5	6
1.1	0.618±0.177	0.534±0.136	0.505±0.215	0.491±0.250	13.65±14.52
1.5	1.539±0.576	0.559±0.137	0.518±0.181	0.586±0.184	0.873±0.451
2.0	0.701±0.085	0.556±0.154	0.427±0.113	0.305±0.104	0.343±0.160
2.5	0.592±0.126	0.484±0.047	0.495±0.122	0.350±0.082	0.495±0.259
3.0	0.533±0.105	0.439±0.098	0.444±0.091	0.406±0.156	0.343±0.100
3.5	0.505±0.084	0.493±0.128	0.417±0.081	0.458±0.183	0.384±0.135
4.0	0.451±0.136	0.454±0.096	0.436±0.068	0.429±0.166	0.444±0.106
4.5	0.498±0.142	0.487±0.084	0.454±0.089	0.443±0.181	0.388±0.044
5.0	0.446±0.092	0.437±0.124	0.413±0.070	0.443±0.094	0.388±0.100



(a) 각 context에 대한 클러스터의 초기 수 : 2



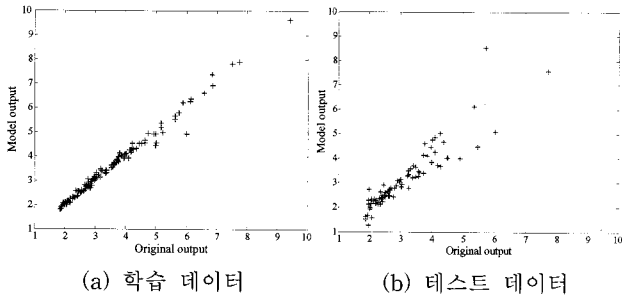
(b) 각 context에 대한 클러스터의 초기 수 : 3

그림 4 Context의 수에 따른 성능지수

Fig. 4 Performance index regarded as a function of the number of contexts

그림 4는 퍼지 계수(m)이 2.0인 경우에 context의 수에 따른 성능 지수(평균값과 표준 오차) 값을 나타낸 것이다.

그림 5는 퍼지 계수(m)가 2.0이고 context의 수는 6 그리고 각각의 context당 초기 클러스터의 수가 2개인 경우, 네트워크의 근사화와 일반화 능력에 대한 선형성을 나타낸 것이다(PI=0.0293, EPI=0.2327).



(a) 학습 데이터

(b) 테스트 데이터

그림 5 모델의 출려과 원 출려과의 차이

Fig. 5 Scatter plots of model output vs. original output (P=6, c=3, m=2.0)

그림 6은 표 1에서 보여준 성능지수 중에서 각각의 퍼지 계수(m)에 따른 최적 네트워크의 성능 지수들의 평균값을 나타낸 것이다. 이는 1.1~5.0 사이에서의 퍼지 계수(m)에 따른 최적값들의 변화를 알 수가 있다.

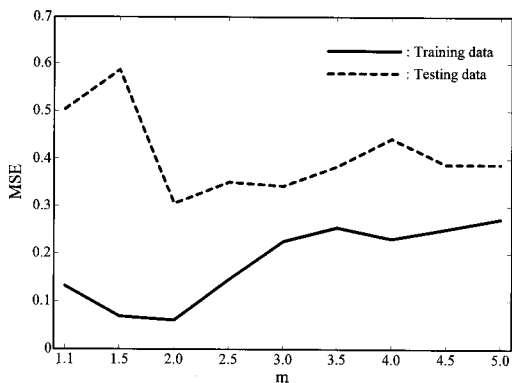


그림 6 퍼지 계수(m)에 따른 최적 네트워크의 성능 지수

Fig. 6 Performance index of the optimized network for different values of the fuzzification coefficient(m)

표 2는 제안된 모델의 성능과 linear regression model 및 표준적인 RBF 뉴럴 네트워크의 성능을 비교한 것이다. 먼저 RBFNN I은 은닉층에서의 가우시안 활성화함수의 중심값을 FCM 클러스터링 방법에 의해서 결정하고 후반부의 계수를 표준적인 최소 자승법을 이용하여 계수를 산출하는 표준적인 RBF 뉴럴 네트워크이다. 또한 은닉층에서의 가우시안 활성화함수의 폭은 '1'로 고정시켰다. RBFNN II 모델 역시 표준적인 RBF 뉴럴 네트워크로써 은닉층의 활성화함수는 FCM 클러스터링 알고리즘을 사용한 모델이다. 출력층에서의 후반부 계수의 추정은 RBFNN I과 같은 최소자승법을 이용한다.

표 2 선택된 모델의 성능과의 비교 분석

Table 2 Comparative analysis of the performance of selected model

Model	No. of nodes in the hidden layer(H)	PI	EPI
Linear regression		0.451±0.091	0.495±0.151
RBFNN I	H=20	2.212±0.371	3.197±1.063
	H=25	2.503±0.595	3.696±1.032
RBFNN II	H=20	0.576±0.109	0.761±0.281
	H=25	0.660±0.209	0.785±0.375
Proposed Model	m=2.0	H≤20	0.059±0.025
	m=2.5	H≤20	0.144±0.058
	m=3.0	H≤24	0.225±0.067
			0.305±0.104
			0.350±0.082
			0.343±0.100

표 2에서 보이듯이, 은닉층에서의 노드 수와 관련하여 제안된 네트워크와의 능력을 비교하기 위해서 본 논문에서 제안된 네트워크에서의 은닉층의 노드의 수와 비슷하도록 노드의 수를 20개와 25개로 구성하였다. 표 2에서 알 수 있듯이 제안된 알고리즘에 의해서 생성된 모델이 기존의 표준적인 RBF 뉴럴 네트워크에 비해서 근사화 능력 및 일반화 능력이 아주 우수함을 알 수 있다.

5.2 Machine Learning Data

5.2.1 가스 터빈 발전소의 NOx 배출 메카니즘 Data

최근 환경보전의 필요성을 크게 인식하며 CO, CO2, SO, SO2, NO, N2O, NO2 및 프레온 가스 등과 같은 공해물질을 줄이고, 이들을 관리하기 위한 연구가 선진국을 중심으로 이루어지고 있다. 본 논문 역시 환경보전의 필요성을 인식하며, 이에 따른 화력발전소 대기오염물질의 배출패턴에 관한 모델을 정립하고 공기 오염물질인 NOx의 배출량을 예측하고자 한다. 특히, 화력발전소의 가스 터빈에서 발생하는 독성이 강한 NOx의 배출 메카니즘(mechanism)에 제안된 알고리즘을 도입하여 배출 패턴 모델을 구축하고 평가한다. 총 5입력(Tamb, COM, LPT, Pcd, Texh) 1출력(NOx)로 구성된 총 260개의 데이터 중, 학습 데이터와 테스트 데이터를 60%-40%로 랜덤하게 나누어 제안된 모델의 성능을 평가하였다.

표 3은 context들의 수와 관련된 최적의 성능을 나타낸 것이다. 표에서도 알 수 있듯이, NOx 데이터의 경우 context의 수가 증가할수록 성능이 향상됨을 알 수 있다.

표 3 제안된 모델의 성능지수

Table 3 Performance index of the proposed model

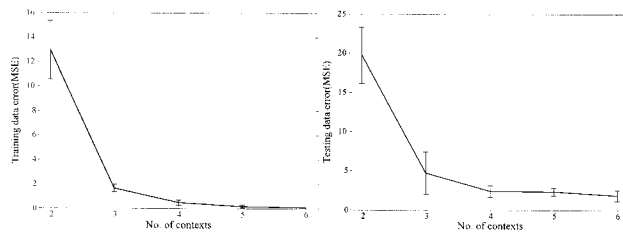
(a) Training data

m	Number of contexts(P)				
	2	3	4	5	6
1.1	225.9±16.08	163.1±29.16	22.69±4.274	11.90±4.898	5.517±2.876
1.5	110.7±28.97	10.89±3.135	4.692±1.305	2.089±1.091	0.300±0.188
2.0	10.26±3.291	0.939±0.309	0.455±0.199	0.164±0.082	0.047±0.015
2.5	31.97±2.783	9.207±1.554	2.522±1.338	0.428±0.226	0.477±0.307
3.0	25.22±1.714	9.955±2.077	7.305±0.928	4.457±0.682	2.312±0.429
3.5	23.81±1.729	9.591±1.692	7.269±0.701	4.444±0.817	2.302±0.484
4.0	21.44±2.041	9.248±1.800	7.584±0.886	4.103±0.690	2.304±0.442
4.5	19.91±1.609	8.639±1.831	7.804±0.965	4.771±0.327	2.278±0.370
5.0	20.05±1.540	9.689±1.422	7.931±0.759	4.361±0.819	2.343±0.309

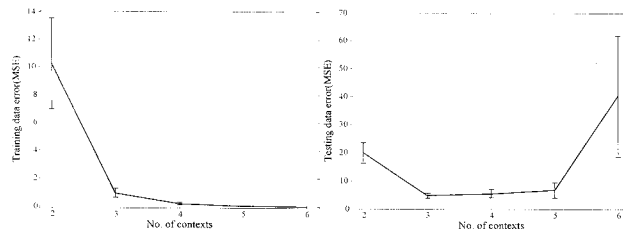
(b) Testing data

m	Number of contexts(P)				
	2	3	4	5	6
1.1	246.2±65.47	381.1±223.9	2695±3825	1.7e5±2.4e5	9.9e4±4.8e4
1.5	109.7±12.33	34.38±14.50	38.88±19.28	31.22±23.27	11.88±6.199
2.0	19.88±3.660	4.787±0.828	2.391±0.729	2.351±0.484	1.831±0.681
2.5	33.84±2.406	11.46±2.142	5.156±1.582	2.177±0.377	1.958±0.509
3.0	28.40±1.731	11.52±0.707	7.988±0.607	5.812±1.234	3.384±0.606
3.5	23.95±0.766	11.58±0.994	9.086±0.632	6.029±0.930	3.695±0.685
4.0	23.07±1.312	11.87±0.751	8.596±0.812	5.800±0.750	3.933±0.450
4.5	22.81±1.674	11.47±1.069	8.588±0.855	6.531±0.753	3.868±0.604
5.0	21.02±1.320	10.63±1.529	9.597±0.929	6.628±0.896	4.208±1.045

그림 7는 퍼지 계수가 2.0인 경우 context의 수에 따른 평균과 표본오차를 나타낸 것이다.



(a) 각 context에 대한 클러스터의 초기 수 : 2



(b) 각 context에 대한 클러스터의 초기 수 : 3

그림 7 Context의 수에 따른 성능지수

Fig. 7 Performance index treated as a function of the number of contexts

그림 8은 퍼지 계수가 2.0일 때 context의 수는 6개이고 각각의 초기 클러스터의 수는 2일 때의 원 데이터의 출력과 모델의 출력과의 차이를 나타낸 그림이다(PI=0.0717, EPI=1.0675).

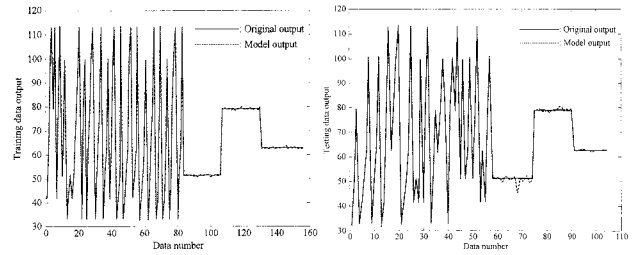


그림 8 원 데이터의 출력과 모델의 출력과의 차이

Fig. 8 Original output versus model output (P=6, c=2, m=2.0)

그림 9에서는 퍼지 계수(m)에 따른 모델의 근사화 능력과 일반화 능력의 성능 차이를 보여준다.

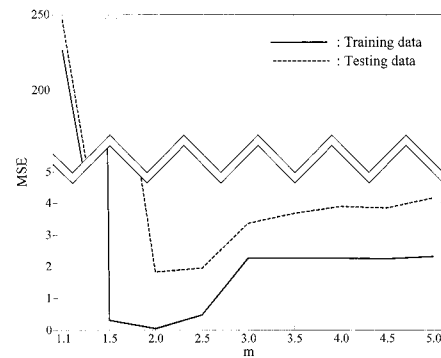


그림 9 퍼지 계수의 서로 다른 값에 따른 네트워크의 성능지수

Fig. 9 Performance index of the network for different values of the fuzzification coefficient

표 4는 RBFNN I, RBFNN II 그리고 선형 회귀 모델과의 제안된 모델과의 성능을 비교한 것이다.

표 4 선택된 모델의 성능과의 비교 분석

Table 4 Comparative analysis of the performance of selected model

Model	No. of nodes in the hidden layer(H)	PI	EPI	
Linear regression		17.81±0.724	18.87±1.075	
RBFNN I	H=20	4.8e3±177.4	4.9e3±256.3	
	H=25	4.8e3±194.7	4.9e3±289.6	
RBFNN II	H=20	184.4±23.56	246.6±45.91	
	H=25	146.9±27.71	241.3±47.46	
Proposed Model	m=2.0	H≤24	0.047±0.015	1.831±0.681
	m=2.5	H≤24	0.477±0.307	1.958±0.509
	m=3.0	H≤24	2.312±0.429	3.384±0.606

5.2.2 Concrete Compressive Strength (CCS) Data

본 실험적 예제에 사용된 CCS 데이터 집합은 콘크리트의 입력으로 수명과 성분(-cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, 그리고 fine aggregate)으로 구성된 매우 비선형적인 함수 형태로 구성

되며 콘크리트 압축 강도가 출력으로 설정된다. 사용된 CCS 데이터 집합은 8입력 1출력 구조의 1030개의 패턴으로 구성되며 앞선 실험적 예제와 마찬가지로 학습 데이터와 테스트 데이터를 랜덤하게 60:40으로 분할하여 평가하였다.

표 5는 제안된 모델의 각각의 context에 대한 최적의 성능지수를 나타낸 것이다. 표 5에서도 알 수 있듯이, 퍼지 계수가 2.0 부근에서 학습 및 테스트에 대한 성능이 우수함을 알 수 있고, 또한 context의 수가 증가할수록 성능의 개선 정도가 뚜렷함을 알 수 있었다.

표 5 제안된 모델의 성능지수

Table 5 Performance index of the proposed model

(a) Training data

m	Number of contexts(P)				
	2	3	4	5	6
1.1	138.3±18.12	68.248±9.48	48.751±9.32	39.952±4.88	32.858±3.21
1.5	122.2±9.542	45.74±5.200	30.19±4.149	33.49±3.839	28.72±2.031
2.0	61.875±2.61	46.218±6.14	33.250±3.21	23.247±2.10	28.237±3.48
2.5	95.009±4.81	92.281±4.82	68.779±4.68	68.540±4.35	55.467±3.41
3.0	96.881±3.69	93.911±3.63	69.028±3.87	66.802±1.97	52.940±3.79
3.5	97.077±3.11	94.155±3.53	69.121±2.77	64.739±2.79	53.991±3.30
4.0	97.368±3.16	93.452±3.89	68.086±4.23	62.587±4.27	52.676±4.03
4.5	97.349±4.06	94.847±4.68	68.278±3.41	59.162±3.05	53.391±4.87
5.0	98.511±5.30	92.474±2.28	68.294±3.01	60.579±4.75	51.901±3.36

(b) Testing data

m	Number of contexts(P)				
	2	3	4	5	6
1.1	208.3±59.14	100.1±17.01	141.9±52.48	360.7±227.6	408.0±173.5
1.5	148.6±12.55	76.673±7.99	77.818±10.3	78.996±15.1	80.203±11.0
2.0	79.291±6.37	79.792±9.35	68.336±7.08	71.670±10.7	67.092±12.9
2.5	103.4±6.771	101.5±6.843	74.743±6.05	80.078±9.90	64.530±5.16
3.0	100.1±4.695	101.1±5.767	70.552±7.75	73.530±5.95	67.550±5.30
3.5	102.9±6.159	100.5±6.567	76.348±7.15	70.419±10.3	66.813±7.33
4.0	101.5±5.750	100.2±5.742	78.960±5.37	73.094±4.66	66.054±4.12
4.5	101.4±5.579	98.753±5.85	76.538±4.93	60.786±4.32	64.342±8.01
5.0	100.3±9.388	101.8±2.890	76.240±7.77	69.406±6.44	63.482±6.78

그림 10은 퍼지 계수가 2.0일 때, 초기 클러스터의 수와 context의 수 증가에 따른 모델의 성능 변화를 나타낸다. Context의 수가 증가할수록 학습 데이터의 성능은 개선되는 것에 반해 테스트 데이터에 대한 성능은 context의 수가 4개일 경우 가장 우수한 성능을 가짐을 알 수 있다.

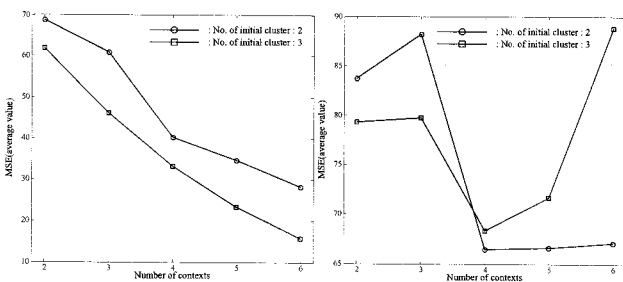


그림 10 Context의 수 증가에 따른 성능지수
Fig. 10 Performance index vis-à-vis number of contexts

표 6은 본 논문에서 제안한 알고리즘의 평가를 위해서 일반적인 형태를 가진 RBFNN I 과 RBFNN II 모델과 더불어 표준적인 모델인 선형 회귀 방정식을 통한 결과를 비교한 것이다.

표 6 다른 모델들과 제안된 모델과의 성능 비교

Table 6 Comparison of the proposed model's performance with other model

Model	No. of nodes in the hidden layer(H)	PI	EPI	
Linear regression		105.8±4.969	111.4±7.866	
RBFNN I	H=20	1.55e3±32.3	1.54e3±42.8	
	H=25	1.55e3±37.4	1.54e3±58.2	
RBFNN II	H=20	181.7±9.463	189.5±10.15	
	H=25	175.8±9.661	183.3±11.29	
Proposed Model	m=2.0	H≤24	28.237±3.48	67.092±12.9
	m=3.5	H≤24	53.991±3.30	66.813±7.33
	m=5.0	H≤24	51.901±3.36	63.482±6.78

6. 결론

본 논문에서는 context-based Fuzzy C-Means 클러스터링을 이용하여 정보 입자화의 세심한 공정을 통해 실험적 데이터의 효과적인 사용을 기반으로 한 Granular-based RBF 뉴럴 네트워크의 새로운 구조를 제안하였다. 본 논문에서 사용한 context-based Fuzzy C-Means는 다음과 같이 2 가지의 중요한 특징을 가지는데 첫 번째는 context의 역할을 강조하는 것이다. 즉 출력 공간에서 정의된 퍼지 집합들(즉 context)을 입력 공간의 정보 입자화의 형성에 중요한 역할을 수행한다. 두 번째로는 실험 데이터들의 표현과 관련된 입자화의 효과를 정량화하는 reconstructability criterion을 통하여 각각의 context에 대해서 적합한 클러스터의 수를 선택할 수 있게 도와준다.

본 연구에서는 포괄적인 실험적 예제를 위해서 저차원 합성 데이터와 공통적으로 많이 사용되어지는 Machine Learning 데이터 집합들을 사용하였다. 또한 널리 사용되어진 RBF 구조들과 선형 회귀 방정식을 이용하여 제안된 모델의 성능을 정량적으로 평가하였다.

결과에서도 알 수 있듯이 이전에 연구되어진 일반적인 RBF 뉴럴 네트워크의 구조에서 얻어진 결과보다 본 논문에서 제안한 알고리즘에서의 성능이 우수함을 알 수 있었으며, 또한 퍼지 계수의 변화에 따른 각각의 데이터들의 특징을 추출할 수 있었다. 더 나아가 각각의 context에 대한 다양한 클러스터 수 사용에 대한 효과적인 메카니즘은 정보 입자화의 중요한 분배에 있어서 효과적인 설계 대체라고 간주되어진다.

감사의 글

이 논문은 2008년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 [KRF-2008-359-D00007].

참 고 문 헌

- [1] Y. Tan, J. Wang, and J. M. Zurada, "Nonlinear blind source separation using a radial basis function network," *IEEE Trans. Neural Networks*, Vol. 12, No. 1, pp. 124-134, 2001.
- [2] M. Han and J. Xi, "Efficient clustering of radial basis perceptron neural network for pattern recognition," *Pattern Recognition*, Vol. 37, pp. 2059-2067, 2004.
- [3] J. M. Vilaplana, J. L. P. Molina, and J. L. Coronado, "Hyper RBF model for accurate reaching in redundant robotic systems," *Neurocomputing*, Vol. 61, pp. 495-501, 2004.
- [4] C. C. Lee, P. C. Chung, J. R. Tsai, and C. I. Chang, "Robust radial basis function neural networks," *IEEE Trans. Syst. Man Cybern. Part B*, Vol. 29, No. 6, pp. 674-685, 1999.
- [5] S. J. Lee and C. L. Hou, "An ART-based construction of RBF networks," *IEEE Trans. Neural Networks*, Vol. 13, No. 6, pp. 1308-1321, 2002.
- [6] M. Wallace, N. Tsapatsoulis, and S. Kollias, "Intelligent initialization of resource allocating RBF networks," *Neural Networks*, Vol. 18, pp. 117-122, 2005.
- [7] C. Harpham and C. W. Dawson, "The effect of different basis functions on a radial basis function network for time series prediction: a comparative study," *Neurocomputing*, Vol. 69, pp. 2161-2170, 2006.
- [8] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.
- [9] J. Moody and C. J. Darken, "Fast learning in networks of locally tuned processing units," *Neural Computation*, Vol. 1, pp. 281-294, 1989.
- [10] W. Pedrycz, "Conditional fuzzy clustering in the design of radial basis function neural networks," *IEEE Trans. Neural Networks*, Vol. 9, No. 4, pp. 601-612, 1998.
- [11] W. Pedrycz, M. G. Chun, and G. Succi, "N4: computing with neural receptive fields," *Neurocomputing*, Vol. 55, pp. 383-401, 2003.
- [12] K. Mali and S. Mitra, "Symbolic classification, clustering and fuzzy radial basis function network," *Fuzzy Sets and Systems*, Vol. 152, pp. 553-564, 2005.
- [13] N. B. Karayiannis, "Reformulated radial basis neural networks trained by gradient descent," *IEEE Trans. Neural Networks*, Vol. 10, No. 3, pp. 657-671, 1999.
- [14] S. M. Loone and G. Irwin, "Improving neural network training solutions using regularization," *Neurocomputing*, Vol. 37, pp. 71-90, 2001.
- [15] H. S. Park, W. Pedrycz, and S. K. Oh, "Evolutionary design of hybrid self-organizing fuzzy polynomial neural networks with the aid of information granulation," *Expert Systems with Applications*, Vol. 33, pp. 830-846, 2007.
- [16] S. K. Oh, W. Pedrycz, and H. S. Park, "Genetically optimized fuzzy polynomial neural networks," *IEEE Trans. fuzzy systems*, Vol. 14, No. 1, pp. 125-144, 2006.
- [17] W. Pedrycz, "Conditional fuzzy c-means," *Pattern Recognition Letters*, Vol. 17, pp. 625-631, 1996.
- [18] W. Pedrycz and K. C. Kwak, "The development of incremental models," *IEEE Trans. Fuzzy Systems*, Vol. 15, No. 3, pp. 507-518, 2007.

저 자 소 개



박 호 성 (朴 浩 成)

1999년 원광대학교 제어계측공학과 졸업. 동 대학원 석사(2001), 박사(2005). 2005-2006년 원광대학교 전기전자 및 정보공학부 전임강사, 2006-2007년 캐나다 Alberta 대학 전기 및 컴퓨터공학과 Post-Doc. 2008~현재 수원대학교 산업기술연구소 학술연구교수. 관심분야는 퍼지추론시스템, 뉴럴 네트워크, 퍼지-뉴럴네트워크, 진화론적 최적화 알고리즘, Granular computing, Neuro-fuzzy computing, Evolutionary computing 등의 Computational Intelligence 등.
Tel : 031-222-6544
Fax : 031-220-2667
E-mail : parkhs@suwon.ac.kr



오 성 권 (吳 聖 權)

1981년 연세대학교 전기공학과 졸업, 동 대학원 석사(1983), 박사(1993). 1983-1989년 금성산전연구소(선임연구원). 1996-1997년 캐나다 Manitoba 대학 전기 및 컴퓨터공학과 Post-Doc. 1993-2004년 원광대학교 전기전자 및 정보공학부 교수. 2005~현재 수원대학교 전기공학과 교수, 2002~현재 대한전기학회, 제어로봇시스템학회, 퍼지및지능시스템학회 편집위원. 관심분야는 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 computational intelligence, 지능 제어 등.
Tel : 031-229-8162
Fax : 031-220-2667
E-mail : ohsk@suwon.ac.kr