입자화기반 RBF 뉴럴네트워크

박호성[1], 오성권[2]
[1]수원대학교 산업기술연구소, [2]수원대학교 공과대학 전기공학과

Granular-based Radial Basis Function Neural Network

Ho-Sung Park[1] and Sung-Kwun Oh[2]
[1]Industry Administration Institute. The University of Suwon
[2]Department of Electrical Engineering. The University of Suwon

Abstract - 본 논문에서는 fuzzy granular computing 방법 중의 하나인 context-based FCM을 이용하여 granular-based radial basis function neural network에 대한 기본적인 개념과 그들의 포괄적인 설계 구조에 대해서 자세히 기술한다. 제안된 모델에 기본이 되는 설계 도구는 context-based fuzzy c-means (C-FCM)로 알려진 fuzzy clustering에 초점이 맞춰져 있으며, 이는 주어진 데이터의 특징에 맞게 공간을 분할함으로써 효율적으로 모델을 구축할 수가 있다.
제안된 모델의 설계 공정은 1) Context fuzzy set에 대한 정의와 설계, 2) Context-based fuzzy clustering에 대한 모델의 적용과 이에 따른 모델 구축의 효율성, 3) 입력과 출력공간에서의 연결된 information granule에 대한 parameter(다항식의 계수들)에 대한 최적화와 같은 단계로 구성되어 있다. Information granule에 대한 parameter들은 성능지수를 최소화하기 위해 Least square method에 의해서 보정된다. 본 논문에서는 모델을 설계함에 있어서 체계적인 설계 알고리즘을 포괄적으로 설명하고 있으며, 더 나아가 제안된 모델의 성능을 다른 표준적인 모델들과 대조함으로써 제안된 모델의 우수성을 나타내고자 한다.

1. 서 론

지능 시스템 또는 computational intelligence 시스템의 최근 연구 경향은 퍼지집합[1], 신경회로망[2] 그리고 진화 알고리즘[3]과 같은 인공지능 도구들의 집적과 K-means 클러스터링 그리고 fuzzy C-means clustering[4]와 관련된 데이터 처리 도구들의 상호 보완적인 합성 형태의 틀 안에서 복잡한 문제를 해결해 나가는 것이다. 퍼지집합, 신경회로망 그리고 진화 알고리즘이 모델링 분야에서 널리 이용되는 반면에, 새로운 방법론적 이슈들을 증가시키고 시스템 모델링에 있어서 많은 문제점을 도출시켰다. 이러한 문제점들을 해결하고자 데이터 전처리 도구로써 K-menas 클러스터링 또는 fuzzy C-menas 클러스터링과 같은 방법들이 computational intelligence 시스템에 적용되어지고 있다. 본 논문에서는 RBFNN(Radial Basis Function Neural Network)의 최적화 문제에 관련된 몇 개의 기존의 문제들을 해결하고자 information granule을 적용하여 접근하고자 한다. 즉, 본 논문은 (1) 주어진 출력 데이터의 특성을 파악하기 위하여 K-means 알고리즘과 삼각형 멤버쉽 함수를 이용하여 context-based FCM[5,6]에 사용되어질 context fuzzy set을 구하고 (2) RBFNN의 receptive field의 수와 출력 데이터를 기반으로 분할된 데이터들의 특징을 입력 데이터에 적용하기 위하여 context-based FCM을 수행하며 (3) 모델의 출력층에 연결된 수학적 함수를 실행하는 것에 목적이 있다. 제안된 RBFNN은 퍼지 집합과 neurocomputing사이의 시너지 효과를 드러내는 neurofuzzy 구조라 할 수 있겠다. 실험적 예제에서도 보이듯이 비선형 데이터들에 대한 제안된 모델의 효율성 및 유연성은 기존의 표준적인 RBF 모델들과 비교하였을 때 그 우수성이 입증됨을 알 수 있다.

2. 본 론

2.1 Context fuzzy set

본 논문에서 사용된 context-based FCM을 이용하기 위해서는 출력 데이터를 먼저 그 특징에 맞게 전처리를 해야 하는 특징을 가지고 있다. 따라서 본 논문에서는 일반적으로 데이터 전처리에 넓게 사용되어진 K-means 알고리즘을 이용하여 주어진 출력데이터를 분류하였다. K-menas 알고리즘에 의해서 구해진 출력 데이터의 prototype을 삼각형 멤버쉽 함수의 정점으로 사용하여 겹칠도가 1/2이 되도록 하여 멤버쉽 값을 얻을 수 있다. 구해진 멤버쉽 값을 이용하여 context-based FCM을 구현하였다.

2.2 Context-based Fuzzy c-means 알고리즘

Context-based FCM은 주어진 출력 데이터의 특징을 해석하여 파악된 출력 데이터의 정보를 기반으로 새로이 입력 데이터를 분할함으로써 기존의 데이터 전처리 방법들보다 더 효율적이고 유기적으로 주어진 입출력 데이터의 특성을 파악하여 모델의 구축에 적용할 수 있다는 장점을 가지고 있다. 이러한 장점을 이용하여 본 논문에서는 context-based FCM 알고리즘을 RBFNN을 구축하는 사용하였다. 단계별 context-based FCM 알고리즘의 수행과정은 다음과 같다.

[단계 1] 각각의 context에 맞게 클러스터의 개수를 결정하고 각각의 context의 소속행렬 U를 다음과 같이 초기화한다.

$$U(f)=\left\{u_{ik}\in[0,1],\ \sum_{i=1}^{c}u_{ik}=f_k\,\forall k,\ 0<\sum_{k=1}^{N}u_{ik}<N\,\forall i\right\} \tag{1}$$

여기서, f_k는 context fuzzy set에서 구한 멤버쉽 함수의 정점을 나타낸다.
[단계 2] [단계 1]로부터 얻어진 u_{ik}값을 기반으로 각각의 클러스터에 대한 중심값과 소속 함수의 값을 구한다.

$$v_i=\{v_{i1},\ v_{i2}\},\quad v_{ij}=\left(\sum_{k=1}^{n}(u_{ik})^m\cdot x\right)/\left(\sum_{k=1}^{n}(u_{ik})^m\right) \tag{2}$$

$$u_{ik}=\frac{f_k}{\sum_{j=1}^{c}\left(\frac{\|x_k-v_i\|}{\|x_k-v_j\|}\right)^{2/(m-1)}} \tag{3}$$

[단계 3] 각각의 클러스터 중심과 데이터와의 거리를 계산하여 새로운 소속행렬을 생성한다.

$$d_{ik}=d(x_k-v_i)=\left[\sum_{j=1}^{l}(x_{kl}-v_{ij})^2\right]^{1/2} \tag{4}$$

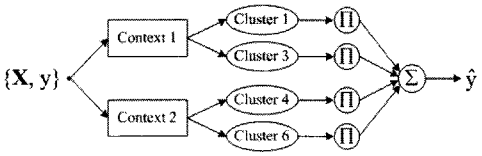
[단계 4] 만일 식 (5)를 만족한다면 종료하고, 그렇지 않으면 r=r+1로 놓고 [단계 2]로 간다.

$$\|U^{(r+1)}-U^{(r)}\|\le\varepsilon(\text{tolerance level}) \tag{5}$$

위의 단계를 각각의 context마다 수행을 하여 각 context에 맞는 클러스터를 구축한다.

2.3 Granular-based RBF 뉴럴네트워크

본 논문에서 제안된 granular-based RBF 뉴럴네트워크의 전체적인 구조는 그림 1과 같이 나타낼 수 있다.



<그림 1> Granular-based BRFNN 구조

그림 1에서 보여준 granular-based RBFNN 구조(context의 수는 2개, 각각의 context에 따른 클러스터의 수는 2개)에서, context는 일반적으로 널리 알려진 K-means 알고리즘을 이용하여 분류하였으며, 분류된 데이터들은 삼각형 멤버쉽 함수를 사용하여 멤버쉽 함수의 값을 구하였다. 각 context에 따른 클러스터의 분할은 본 논문에서 중심적으로 다룬 context-based FCM 알고리즘을 이용하여 각 context에 맞게 분할된 데이터를 기반으로 입력 데이터를 분할하여, 데이터 특성에 맞는 RBFNN을 구축하였다.
모델의 최종 모델식은 다음과 같다.

$$\hat{y}=\sum_{i=1}^{p}\sum_{j=1}^{c}y_{ij}\ ,\quad y_{ij}=u_{ij}\left(\sum_{k=0}^{n}a_{ijk}(x_k)\right) \tag{6}$$

여기서, i는 context의 수, j는 각각의 context에 따른 클러스터의 수, k는 입력 데이터의 수, y_{ij}는 모델의 local 출력식을 의미한다. 식 (6)의 매개변수 a_{ijk}는 표준 최소자승법으로 추정하였다.

3. 실험적 예제

제안된 방법론에 대한 효과적인 분석적 해석을 돕기 위해 잘 알려지고 넓게 사용된 2개의 데이터 집합을 통해서 제안된 모델의 성능을 나타내고자한다. 모델의 성능 지수식으로는 MSE(Mean Squre Error)을 이용하였으며, 모든 실험적 예제들은 10-fold cross-validation를 수행하였다.

3.1 Synthetic One-Dimensional Data

본 논문에서는 식 (7)과 같이 1입력 1출력을 가진 비선형 함수를 이용하여 제안된 알고리즘의 성능을 비교하였다.

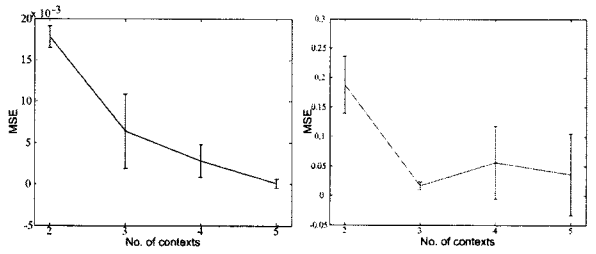
$$y = 0.6\sin(\pi x) + 0.3\sin(3\pi x) + 0.1\sin(5\pi x), \quad X = [-1, 1] \tag{7}$$

식 (7)를 이용하여, 전체 200개의 입출력 데이터 집합을 생성한 후, 학습 데이터(60%)와 테스트 데이터(40%)를 랜덤 분할하였다.
제안된 모델의 다양한 정보의 입자화를 평가하기 위해, 본 논문에서는 context의 수를 2~5개를 사용하였으며, 각각의 context에 대해서 클러스터를 2~6로 분할하여 모델에 적용하였다. 표 1은 제안된 모델의 각각의 context에 대한 성능을 요약한 것이다.

<표 1> RBF 뉴럴 네트워크의 성능지수

Number of clusters	Number of contexts							
	2		3		4		5	
	PI	EPI	PI	EPI	PI	EPI	PI	EPI
2	0.0178 ±0.0013	0.1871 ±0.0483	0.0064 ±0.0045	0.0171 ±0.0065	0.0028 ±0.0020	0.0562 ±0.0616	9.34e-4 ±5.5e-4	0.0362 ±0.0689
3	0.0171 ±0.0025	0.0533 ±0.0139	0.0037 ±0.0014	0.6323 ±0.6652	0.0013 ±5.3e-4	0.6339 ±0.7432	5.25e-4 ±4.0e-4	0.3267 ±0.4943
4	0.0073 ±0.0015	0.5439 ±0.2288	8.60e-4 ±5.5e-4	0.1438 ±0.1356	5.43e-4 ±3.5e-4	0.4397 ±0.5533	2.12e-4 ±1.0e-4	0.2418 ±0.1879
5	0.0057 ±6.3e-4	0.1978 ±0.1123	4.04e-4 ±2.3e-4	1.5854 ±1.1934	2.19e-4 ±1.4e-4	0.2027 ±0.2688	7.59e-5 ±5.9e-5	0.4568 ±0.6334
6	0.0020 ±3.2e-4	0.4142 ±0.2105	1.76e-4 ±7.5e-5	1.2217 ±0.7128	1.25e-4 ±7.6e-5	0.3045 ±0.2881	2.66e-5 ±2.0e-5	0.1937 ±0.2439

그림 2는 클러스터의 수가 2일때 context의 수의 증가에 따른 성능지수의 값들을 평균과 표준오차를 나타낸 것이다.



(a) Training data (b) Testing data

<그림 2> Context의 수에 따른 성능지수

표 1과 그림 2에서 볼 수 있듯이, 제안된 데이터 집합에서는 context의 수가 3이고 이에 따른 클러스터의 수가 2일때 학습데이터와 테스트 데이터의 성능지수를 상호 고려하였을 때 가장 우수함을 알 수 있다.
표 2는 기존의 표준적인 RBFNN과의 성능지수를 비교한 것을 보여준다. 표 2에서도 알 수 있듯이 제안된 모델의 성능이 우수함을 알 수 있다.

<표 2> 선택된 RBFNN 성능의 비교 분석

Model	No. of nodes in the hidden layer	PI	EPI
RBFNN I	5	0.0387 ± 0.0028	0.0402 ± 0.0042
RBFNN II	5	0.0522 ± 0.0050	0.0565 ± 0.0090
Proposed model	6(p=2, c=3)	0.0171 ± 0.0025	0.0533 ± 0.0139
	6(p=3, c=2)	0.0064 ± 0.0045	0.0171 ± 0.0065

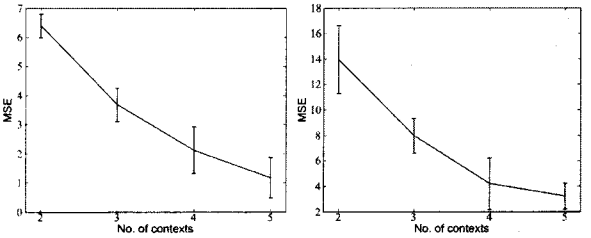
표 2에서 RBFNN I의 모델은 FCM 알고리즘을 이용하여 가우시안 형태의 receptive field 중심값을 구하였으며, 은닉층과 출력층사이의 연결가중치 값은 표준 최소자승법을 이용한 모델이다. RBFNN II 모델은 receptive field의 함수를 FCM 알고리즘 자체를 이용한 것이며, RBFNN I과 마찬가지고 표준 최소자승법을 이용하여 parameter를 추정한 모델이다. 표 2에서 'p'의 표기는 context의 수를 의미하고, 'c'는 각각의 context에 분할된 클러스터의 수를 의미한다.

3.2 Machine learning data

Machine learning repository로부터 얻은 데이터 집합 중에서 'automobile miles per gallon(MPG) data'를 이용하여 본 논문의 성능을 입증한다. 사용된 데이터 집합은 전체 392개의 입출력 데이터 중에서 학습과 테스트 데이터를 60-40%로 랜덤하게 나누어 모델을 구축하였다.
표 2에서 알 수 있듯이, MPG 데이터 집합을 이용한 경우에서는 synthetic one-dimensional data의 성능분포와는 달리 context의 수가 증가 할수록 그 성능은 개선이 되어지지만 클러스터의 수의 증가에 따른 모델의 성능은 크게 영향을 미치지 않음을 알 수 있다.
그림 3은 클러스터의 수가 2일때 context의 수의 증가에 따른 성능지수의 값들을 평균과 표준오차를 나타낸 것이다. 표 2와 그림 3에서도 볼 수 있듯이 주어진 데이터에 대한 평균오차의 정도가 학습 데이터보다는 테스트 데이터에 대해서 심함을 알 수 있으며 이는 모델의 일반화 능력보다는 근사화 능력이 뛰어남을 보여준다.

<표 3> 제안된 모델의 성능지수

Number of clusters	Number of contexts							
	2		3		4		5	
	PI	EPI	PI	EPI	PI	EPI	PI	EPI
2	6.3915 ±0.4134	13.9607 ±2.6761	3.6886 ±0.5750	7.9806 ±1.3722	2.1212 ±0.8038	4.2033 ±2.0414	1.1819 ±0.7007	3.2249 ±1.0219
3	5.6551 ±0.6962	21.6128 ±7.5841	3.0812 ±0.4167	11.5804 ±4.3563	1.7014 ±0.7733	6.9344 ±0.7691	1.2567 ±0.6858	14.8634 ±8.0169
4	5.4724 ±0.5925	19.7506 ±4.4848	2.6117 ±0.4054	29.0389 ±16.116	1.6829 ±0.6843	28.8358 ±22.801	1.0352 ±0.6657	23.6616 ±19.148
5	5.1853 ±0.5183	40.9332 ±15.558	2.3965 ±0.4296	47.7761 ±31.746	1.3148 ±0.665	42.265 ±24.814	0.9196 ±0.7828	42.8895 ±23.279
6	4.7098 ±0.3983	64.9353 ±18.417	2.1312 ±0.4479	123.279 ±283.10	0.9096 ±0.4848	52.6818 ±26.586	0.6195 ±0.7479	310.549 ±251.41



(a) Training data (b) Testing data

<그림 3> Context의 수에 따른 성능지수

표 3은 표준 모델인 RBFNN I과 RBFNN II과의 성능비교를 나타낸 것이다. 표 3에서도 볼 수 있듯이 제안된 RBFNN모델이 다른 표준적인 모델에 비해서 우수함을 보여준다.

<표 3> 표준 RBFNN과의 성능 비교

Model	No. of nodes in the hidden layer	PI	EPI
RBFNN I	10	26.976 ± 3.6016	34.289 ± 8.5850
RBFNN II	10	16.599 ± 0.5350	18.347 ± 1.2820
Proposed model	8(p=4, c=2)	2.1212 ± 0.8038	4.2033 ± 2.0414
	10(p=5, c=2)	1.1819 ± 0.7007	3.2249 ± 1.0219

4. 결 론

본 논문에서는 비선형 데이터를 이용한 효율적인 모델 구축을 위하여 information granularity를 이용한 context-based FCM 방법을 기반한 granular-based RBFNN의 구체적인 설계 절차 및 구조에 대해서 자세히 기술하였다. 일반적인 RBFNN과 비교하였을 때, 제안된 모델은 시스템 구축에 사용된 다양한 비선형 특징을 가진 데이터를 효과적으로 처리하기 위해서 context-based FCM을 사용하여 각각의 context에 비슷한 특성을 가진 데이터를 분할한 다음, 분할된 데이터의 특징을 고려하여 모델을 구축하였다. 이는 실험적 예제에서도 보이듯이, 네트워크의 최종 구조에 비선형 특징을 가진 데이터의 특징을 기여할 뿐만 아니라 유연성 있는 모델을 구축할 수 있다는 장점을 가지고 있다.

감사의 글

본 연구는 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음 [GGA0801-45700, U-city 보안감시 기술협력센터] 그리고 2008년도 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2008-314-D00376)

[참 고 문 헌]

[1] L. A. Zadeh, "Outline of a new approch to the analysis of complex system and decision process", IEEE Tran. on SMC, Vol 1, pp. 28-44, 1973
[2] J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System", IEEE Tran. SMC, Vol. 23, No. 3, pp. 665-685, 1993
[3]D. E. Goldberg, "Genetic Algorithm in search, Optimization & Machine Learning", Addison-wesley, 1989
[4] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective function Algorithm", New York, Plenum, 1981
[5] W. Pedrycz, "Conditional fuzzy C-means", Pattern Recognition Letters, Vol. 17, No. 6, pp. 625-632, 1996
[6] W. Pedrycz, "Conditional fuzzy clustering in the design of radial basis function neural network", IEEE Tran. Neural Network, Vol. 9, No. 4, pp. 601-612, 1998