

차분 진화 알고리즘 기반 방사형 기저 함수 신경회로망 분류기의 최적화 방법

마창민*, 오성권**
수원대학교*

Optimization Method of Differential Evolution-based Radial Basis Function Neural Networks

Chang-Min Ma*, Sung-Kwun Oh**
The University of Suwon*

Abstract - 본 연구에서는 패턴분류를 위해 최적화된 방사형 기저 함수 신경회로망(Radial Basis Function Neural Networks) 분류기를 제안한다. RBFNN은 입력층, 은닉층, 출력층의 3층 구조로 되어 있으며 Multi Dimension, Predictive ability, Robustness한 특징이 있다. RBFNN의 은닉층에는 기존의 활성화함수가 아닌 Fuzzy C-means 클러스터링 알고리즘을 사용하여 입력 데이터의 특성을 고려한 적합도를 사용하였다. RBFNN은 은닉층의 노드수와 FCM 클러스터링의 퍼지화 계수, 연결가중치의 다항식 타입이 모델의 성능의 향상에 영향을 미치기 때문에 최적화가 필요하며 본 논문에서는 Differential Evolution(DE) 알고리즘을 사용하여 모델의 구조 및 파라미터를 최적화시켜 모델의 성능을 향상시켰다. 제안된 모델을 평가하기 위해 패턴분류에 많이 사용되는 Iris 데이터와 Wine 데이터를 이용하였다.

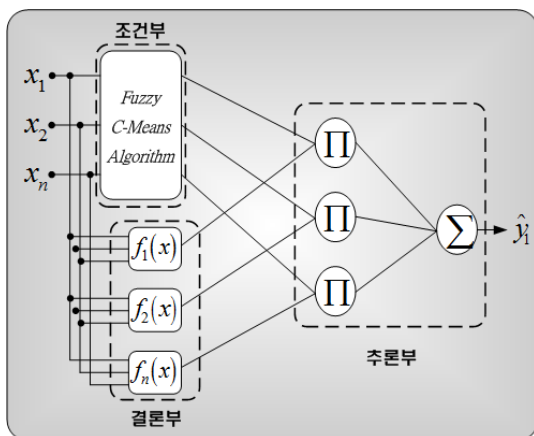
1. 서 론

신경회로망을 기반으로 설계된 분류기는 학습능력과 일반화 능력의 우수성으로 다양한 패턴인식 시스템에서 가장 많이 사용되는 분류기 중 하나이다. 신경회로망을 기반으로 설계된 분류기는 입력층, 다수의 은닉층, 출력층을 갖는 다층구조가 일반적이며 경사 하강법을 이용하여 학습된다. 본 논문에서는 신경회로망을 기반으로 한 방사형 기저 함수 신경회로망(Radial Basis Function Neural Networks; RBFNN)을 설계하여 패턴분류기로 사용하고 최적화 시킨다. 제안하는 모델은 FCM 클러스터링에 기반 한 분할 함수를 은닉층의 활성화함수로 사용하여 입력공간의 데이터 분포 특성을 잘 반영하도록 설계한다. 또한 최적화 알고리즘의 하나인 차분 진화(Differential Evolution; DE) 알고리즘을 이용하여 모델에 가장 알맞은 클러스터의 수와 퍼지화 계수, 연결 가중치의 다항식의 차수를 최적화 시켜 판별식을 찾아낸다. 본 논문에서는 패턴 분류기로써 제안된 모델의 평가를 위해서 Iris 데이터와 Wine 데이터를 이용한다.

2. 본 론

2.1 방사형 기저 함수 신경회로망

기본적인 신경회로망의 구조는 입력층과 은닉층, 출력층으로 구성된다. RBFNN은 기본적인 신경회로망의 구조를 기반으로 입력층과 은닉층, 출력층이 각 1개의 층만 존재하며 Multi Dimension, Predictive ability, Robustness한 특징을 갖고 있다.



〈그림 1〉 제안된 RBFNN 패턴 분류기의 구조

그림 1은 본 논문에서 제안한 RBFNN 패턴 분류기의 구조이다. RBFNN의 조건부는 방사형의 활성화함수를 사용하는 것이 아닌 Fuzzy C-Means 클러스터링 알고리즘을 사용한다. FCM 알고리즘은 각 클러스터에 포함되는 데이터의 소속정도가 퍼지집합으로 출력되어 방사형의 활성화함수 형태를 나타낸다. 결론부에서는 표 1과 같이 상수항을 제외한 일차식, 2차식, 변형된 2차식의 3가지 타입 중 하나의 형태를 갖는다.

〈표 1〉 다항식 차수 (입력이 2개일 경우)

Type	Polynomial Type
Linear	$\hat{y} = a_0 + a_1x_1 + a_2x_2$
Quadratic	$\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_2^2 + a_5x_1x_2$
Modified Quadratic	$\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2$

추론부에서는 네트워크의 최종출력을 구하게 되며 그림 1에서 “Π”로 표기된 뉴런은 조건부의 출력과 결론부의 연결가중치의 곱을 나타내며 “Σ”로 표기된 뉴런에 의해 입력신호들이 합쳐져 그 결과를 출력층 뉴런의 최종 출력으로 내보낸다.

2.2 Fuzzy C-means(FCM) Clustering 알고리즘

클러스터링(Clustering) 방법이란 데이터의 분류를 위해서 사용되는 것으로 데이터의 내부가 비슷한 패턴, 속성, 형태 등의 기준을 통해 데이터를 분류하여 개체의 소속집단을 정의해주는 것이다. 본 논문에서는 클러스터에 대한 각 데이터의 거리를 기준으로 소속정도를 측정하여 데이터들의 특성을 분류하는 FCM 클러스터링 방법을 사용하였다. 다음은 FCM 클러스터링의 절차이다.

[단계 1] 클러스터의 개수를 결정하고, 소속행렬 $U^{(r)}$ 을 초기화한다.

$$U^{(r)} = \left\{ u_{ik} \in [0, 1], \sum_{i=1}^c u_{ik} = 1 \forall k, 0 < \sum_{k=1}^n u_{ik} < n \forall i \right\} \quad (1)$$

[단계 2] $U^{(r)}$ 값을 기반으로 각각의 클러스터에 대한 중심값과 소속함수의 값을 구한다.

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

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

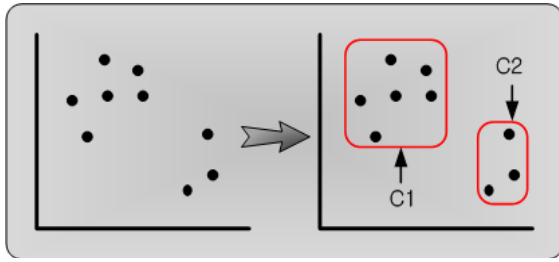
[단계 3] 각각의 클러스터 중심과 데이터와의 거리를 계산하여 새로운 소속행렬을 생성한다.

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

[단계 4] 오차가 허용범위 안에 도달하면 종료하고, 그렇지 않으면 단계 2로 돌아간다.

$$\| U^{(r+1)} - U^{(r)} \| \leq \epsilon (\text{tolerance level})$$

데이터를 FCM 클러스터링 알고리즘으로 분류하고 나오는 소속함수의 형태가 방사형을 이루고 있어 RBFNN의 은닉층의 활성화함수 대신 FCM 클러스터링의 소속함수를 적용하였다.



〈그림 2〉 FCM 클러스터링 방법의 예

2.3 차분 진화(DE) 알고리즘을 이용한 최적화 방법

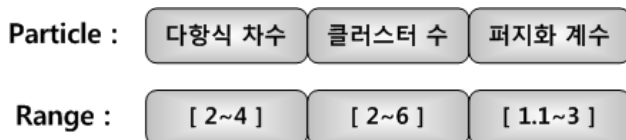
차분 진화(Differential Evolution; DE) 알고리즘은 Price와 Storn에 의해 벡터 차분을 사용하여 Chebychev 다항곡선의 내삽문제를 해결하는 과정에서 개발되었다. DE는 간단한 구조를 갖으면서도 전역 최적화에 대한 수렴성이 뛰어나고 다른 진화 알고리즘들에 비해서 컴퓨팅 시간이 짧은 장점을 지닌다. DE는 해집단 기반의 병렬탐색을 수행하며 다른 진화 알고리즘과는 달리 차분 진화는 확률분포에 기반한 변이이산자(mutation operator)에 의존하지 않는다. 차분진화에서 사용되는 연산자는 주로 임의로 선택된 개체들 사이의 차이를 사용한다. 일반적으로 실수값을 갖는 문제해결에 사용되는 차분진화는 두 벡터의 차이와 또 다른 벡터와 가중합을 계산한다. 다음은 차분진화 알고리즘의 절차이다.

- [단계 1] 초기 집단 구성(랜덤 값으로 μ 개의 개체를 초기화. 각 개체는 n 개의 목적변수로 구성, $t=0$)
- [단계 2] 목표벡터의 임의선택, (x_1, x_2, x_3)
- [단계 3] 가중된 차이벡터의 생성, $r(x_1 - x_2)$
- [단계 4] x_3 를 $F(x_1 - x_2)$ 와 더하여 시행벡터 v 를 생성,

$$v = x_3 + F(x_1 - x_2) \quad (2)$$
- [단계 5] 시행벡터와 현재벡터 x_i 와 교차비(CR, crossover rate)에 의해 교차연산 수행
- [단계 6] x_i 의 교체 또는 유지

위의 절차에서 초기에 해집단을 생성한다. 초기에 결정된 집단의 크기는 진화과정에서 변화하지 않는다. 해집단으로부터 3개의 목표벡터(target vector)를 임의추출한다. 차분진화는 두 개의 개체벡터의 차이에 가중치 r 를 곱한 것을 나머지 한 개의 개체 벡터에 더해서 교차용 시행벡터(trial vector)를 생성한다. 시행벡터는 현재벡터와 교차비에 의해 교차연산을 수행한다. 최종적으로 x_i 의 대체 혹은 유지를 결정한다.

DE를 이용하여 제안된 모델의 입력변수의 수와 다항식의 차수, 클러스터의 수 및 퍼지화 계수를 최적화 시킨다. 최적화를 위해 필요한 개체의 구조는 그림 3과 같이 나타낼 수 있다.



〈그림 3〉 차분 진화 알고리즘의 개체 구조

총 3개의 파라미터를 최적화 시키며 입력변수의 경우 2개부터 주어진 데이터의 최대 입력개수 안에서 선택된다. 다항식 타입은 표 1과 같이 4개의 타입 중에서 선택되고 클러스터의 수는 2개에서 5개 사이로 설정하였다.

3. 결 론

제안된 모델의 성능을 평가하기 위해 패턴분류에서 많이 사용되는 Iris 데이터와 Wine 데이터를 사용하였다. Iris dataset은 총 4입력 3클래스로 Iris라는 꽃을 3종류로 나눈 데이터이고 Wine dataset은 총 13입력 3클래스로 구성되어 있다. 데이터는 각 클래스를 6:4로 학습 데이터와 테스트 데이터로 나누어 실험하고 결과는 테스트 데이터의 분류률로 나타내었다. 제안된 모델의 최적화를 위한 DE의 파라미터 및 탐색범위는 표 2와 같다.

표 3은 Iris dataset을 패턴분류한 결과이다. 표 3에 있는 패턴의 분류률을 보면 RBFNN의 은닉층 노드의 수, 본 논문에서는 FCM 알고리즘의 클러스터의 수가 2개 이고 연결가중치의 다항식 타입이 일차식과 2차식일 때, 클러스터의 수가 6개 이고 연결가중치의 다항식 타입이 일차식과 2차식인 경우 100%의 분류률을 보였고 전체적으로도 90% 이상의

높은 분류률을 나타내었다.

〈표 2〉 DE 파라미터 및 탐색범위

Parameters		Value
Number of generations		100
Number of population		50
Crossover rate		0.8
F		0.95
The range of search space	Polynomial Type	[2~4]
	Number of clusters	[2~6]
	Fuzzification coefficient	[1.1~3]

〈표 3〉 Iris dataset의 다항식 타입과 를 수에 따른 분류률

	Linear	Quadratic	Modified Quadratic
2	100.00 %	100.00 %	98.48 %
3	98.48 %	98.48 %	98.48 %
4	98.48 %	98.48 %	96.97 %
5	98.48 %	95.45 %	98.48 %
6	100.00 %	100.00 %	95.45 %

〈표 4〉 Wine dataset의 다항식 타입과 를 수에 따른 분류률

	Linear	Quadratic	Modified Quadratic
2	98.77 %	90.12 %	82.72 %
3	100.00 %	88.89 %	93.83 %
4	100.00 %	92.59 %	88.89 %
5	98.77 %	91.36 %	92.59 %
6	98.77 %	82.72 %	91.36 %

표 4는 Wine dataset을 패턴 분류한 결과로 다항식의 타입이 일차식이고 클러스터의 수가 3개이거나 4개인 경우 100%의 분류률을 나타내었다.

본 논문에서는 제안된 RBFNN 분류기의 최적화를 통하여 패턴 분류 데이터들을 분류하고 각 조건에 따른 분류률을 나타내었다. 최적화를 통하여 가장 좋은 성능을 나타내는 분류기의 조건과 제안된 분류기의 성능이 우수함을 알 수 있었다.

감사의 글

본 연구는 2009년도 지식경제부의 재원으로 한국에너지 기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다(2009T100100563) 그리고 이 논문은 2010년도 정부(과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2010-D00065).

[참 고 문 헌]

- [1] M. L. Kothari, S. Madhani, and R. Segal, "Orthogonal Least Square Learning Algorithm Based Radial Basis Function Network Adaptive Power System Stabilizer" Proceeding of IEEE SMC, Vol. 1, pp. 542-547, 1997.
- [2] K. Z. Mao, "RBF Neural Network Center Selection Based on Fisher Ratio Class Separability Measure", IEEE Trans. Neural Network, Vol 13(5), pp.1211-1217, 2002.
- [3] R. Storn, K. V. Price, "Differential Evolution—a fast and efficient heuristic for global optimization over continuous spaces", Journal of Global Optimization, vol. 11, pp. 341-359, 1997.
- [4] 박진열, 박병준, 오성권, "퍼지추론 기반 다항식 RBF 뉴럴 네트워크의 설계 및 최적화," 전기학회논문지, Vol.58, No. 2, pp.339-406, 2009