ISSN 1975-8359(Print) / ISSN 2287-4364(Online)
The Transactions of the Korean Institute of Electrical Engineers Vol. 67, No. 8, pp. 1071~1079, 2018
http://doi.org/10.5370/KIEE.2018.67.8.1071

입자군집 최적화를 이용한 SVM 기반 다항식 뉴럴 네트워크 분류기 설계

Design of SVM-Based Polynomial Neural Networks Classifier Using Particle Swarm Optimization

노 석 범* · 오 성 권†
(Seok-Beom Roh · Sung-Kwun Oh)

Abstract - In this study, the design methodology as well as network architecture of Support Vector Machine based Polynomial Neural Network, which is a kind of the dynamically generated neural networks, is introduced. The Support Vector Machine based polynomial neural networks is given as a novel network architecture redesigned with the aid of polynomial neural networks and Support Vector Machine. The generic polynomial neural networks, whose nodes are made of polynomials, are dynamically generated in each layer-wise. The individual nodes of the support vector machine based polynomial neural networks is constructed as a support vector machine, and the nodes as well as layers of the support vector machine based polynomial neural networks are dynamically generated as like the generation process of the generic polynomial neural networks. Support vector machine is well known as a sort of robust pattern classifiers. In addition, in order to enhance the structural flexibility as well as the classification performance of the proposed classifier, multi-objective particle swarm optimization is used. In other words, the optimization algorithm leads to sequentially successive generation of each layer of support vector based polynomial neural networks. The bench mark data sets are used to demonstrate the pattern classification performance of the proposed classifiers through the comparison of the generalization ability of the proposed classifier with some already studied classifiers.

Key Words : Support vector machine, Polynomial neural networks, Dynamically generated networks, Support vector machine based polynomial neural networks, Multi-objective particle swarm optimization

1. 서 론

오늘날 현재에는 개인들이 소유하고 있는 다양한 모바일 통신 기기들의 발달과 IT 기술의 발달로 인하여 개인사용자들이 생산하고 배포하는 텍스트 및 이미지와 같은 다양한 형태의 데이터들의 생산량이 급증하고 있다. 이와 같이 개인들의 요구와 상황을 대변할 수 있는 데이터를 처리하여 대용량의 데이터로부터 유의미한 정보를 획득할 수 있는 기술을 개발하기 위한 다양한 연구가 진행되고 있다.

이러한 데이터로부터 유의미한 정보를 추출하기 위한 연구 중에서 주어진 데이터를 분류하기 위한 패턴 분류기에 대한 연구가 활발히 진행되고 있다. 패턴 분류 기법은 매우 다양한 관점에서 접근이 가능하고, 다양한 관점에서 기술이 개발되어지고 있다. 패턴 분류 기법은 크게 분류하면, 통계적 접근 방법과 뉴럴 네트워크를 이용한 방법으로 나눠질 수 있다 [1].

또한 패턴 분류기 설계의 경우 입력 공간의 차원이 매우 큰 경우 이와 같은 데이터를 정확히 분류하는 것은 어려운 문제로 알려져 있다. 일반적으로 고차원의 입력 공간상에서 패턴 분류하기 위하여 주성분 분석(Principle Component Analysis; PCA) 또는 선형 판별식 분석(Linear Discriminant Analysis; LDA)와 같은 차원 축소 알고리즘을 이용하여 데이터 입력 공간을 축소한 후 다양한 패턴 분류 알고리즘을 적용한다. 또 다른 연구방향으로 차원 축소 알고리즘을 적용하지 않고 네트워크의 생성 특성을 이용한 다항식 뉴럴 네트워크(Polynomial Neural Networks; PNNs)를 이용한 고차원 데이터 문제 해결 방법이 있다.

본 논문에서는 고차원 데이터의 패턴 분류 성능을 개선하기 위하여 동적 생성이 가능한 다항식 뉴럴 네트워크의 설계 방법을 기반으로 패턴 분류기를 설계하고자 한다. 일반적으로 신경회로망 기반 패턴 분류기는 입력 데이터의 전체 변수를 입력으로 사용하고 네트워크의 전체구조가 파라미터 학습을 진행하기 전에 미리 결정되어 진다. 반면 다항식 뉴럴 네트워크의 뉴런들은 입력 변수들 중에서 설계자가 결정한 개수의 입력 변수만 입력으로 사용한다. 또한 네트워크의 생성 방식 측면에서도 전체 네트워크의 구조가 미리 결정되는 기존 신경회로망과 달리 다항식 뉴럴

† Corresponding Author : Dept. of Electrical and Electronic Engineering, University of Suwon, Korea.
E-mail: ohsk@suwon.ac.kr
* Dept. of Electrical and Electronic Engineering, University of Suwon, Korea.
Received : April 6, 2018; Accepted : May 1, 2018

Copyright ⓒ The Korean Institute of Electrical Engineers

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/)which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

네트워크는 동적으로 한 층 한 층 생성되어진다. 다항식 뉴럴 네트워크의 동적 생성 특성으로 인하여 매우 유연한 네트워크의 구조를 얻을 수 있다.

일반적인 신경회로망의 경우 오류 역전파 알고리즘 (Error Back-Propagation) 알고리즘과 같은 반복적인 학습을 통하여 네트워크의 파라미터를 최적화 하는 반면, 다항식 신경회로망은 각 층의 기본 구조인 다항식 뉴런의 (Polynomial Neuron)의 계수를 추정하기 위하여 Least Square Estimation 방법을 사용하여 빠른 학습 속도를 보인다. 그러나 이와 같이 고차원 문제에 적합하고 빠른 학습속도를 가지는 다항식 뉴럴 네트워크도 네트워크의 복잡도를 증가시키기 위하여 네트워크의 층수를 증가시키면, 다항식 뉴럴 네트워크 패턴 분류기가 학습데이터에 과적합 (overfitting) 되는 특성을 보인다[2]. 과적한 문제를 해결하기 위하여 매우 다양한 방법들의 제안되어져 왔고 지금도 연구되어지고 있다. 본 논문에서는 동적생성 다항식 뉴럴 네트워크의 층수를 증가 시킬 때 나타나는 과적합 문제를 감소시키기 위하여 강인한 패턴 분류기로 알려진 Support Vector Machine (SVM)을 신경회로망의 기본 구조인 다항식 뉴런으로 사용한다. SVM은 1990년대 초반에 Vapnik에 의해 제안된 방법으로 패턴 분류 및 모델링 분야에 적용되어 우수한 성과를 거두고 있다[2]. 또한 SVM은 얼굴인식, 고장진단과 같은 다양한 분야에 많이 이용되어 우수한 성능을 보이고 있다[3-7].

이와 같이 우수한 패턴 분류 성능을 보이는 SVM을 다항식 신경회로망의 기본 노드로 사용함으로써, 동적 생성 네트워크가 가지는 문제점인 과적합 문제를 해결하고자 한다.

제안된 SVM 기반 다항식 뉴럴 네트워크의 최적화를 위하여 다중 목적 입자 군집 최적화 (Multi-Objective Particle Swarm Optimization: MOPSO) 알고리즘을 사용하였다 [5]. 일반적으로 하나의 목적함수를 가지는 입자 군집 최적화 알고리즘과 달리 다수의 목적함수를 기반으로 패턴 분류기를 최적화시키기 위하여 다중 목적 입자 군집 최적화 알고리즘을 이용한다.

제안된 다항식 뉴럴 네트워크의 패턴 분류 성능을 평가하기 위하여 다양한 기계 학습 데이터들 중에서 binary class 분류 데이터들을 적용하여 평가한다.

본 논문의 구성은 다음과 같다. 2장에서는 동적 생성 신경회로망인 다항식 뉴럴 네트워크에 대하여 설명하고, 3장에서는 Support Vector Machine 에 대해서 알아보고 SVM 기반 다항시 뉴럴 네트워크를 제안한다. 4장에서는 다중 목적 입자 군집 최적화 알고리즘을 이용한 제안된 SVM 기반 다항식 뉴럴 네트워크의 최적화에 대해 설명하고 5장에서는 제안된 패턴 분류기의 성능을 평가하기 위해 Benchmark 데이터를 이용하여 얻은 실험 결과를 제시하고 6장에서는 결론을 도출한다.

2. 다항식 뉴럴 네트워크

다항식 뉴럴 네트워크는 1968년 A. G. Ivakhnenko에 의해 제안된 Group Method of Data Handling (GMDH) 모델을 확장한 형태이며, 동적으로 생성되는 특징을 가진 네트워크이다[2]. 일반적으로 전체 입력 변수를 사용하는 신경회로망과는 달리, 다항식 뉴럴 네트워크의 노드는 전체 입력 변수들 중에서 설계자가 미리 정한 개수의 입력변수만을 사용한다. 또한 네트워크의 구조가 미리 결정되고, 결정된 구조를 구성하고 있는 연결 하중값을 학습 알고리즘을 통해 결정하는 일반적인 신경회로망과 달리 다항식 뉴럴 네트워크의 경우, 네트워크의 각 층이 동적으로 생성되어 네트워크의 전체구조가 모든 학습 과정을 거친 후에 결정되는 특징을 가지고 있다. 따라서 다항식 뉴널 네트워크는 기존 신경회로망에 비해 적은 노드수와 유연한 적응 특성을 보인다 [3-6].

다항식 뉴럴 네트워크의 구조는 그림 1과 같다.

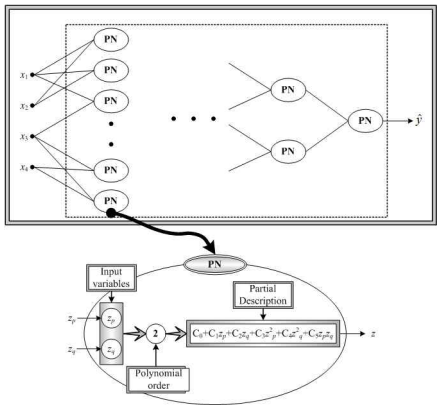


그림 1 다항식 뉴럴 네트워크의 구조

Fig. 1 Architecture of generic polynomial neural networks

그림 1에서 PN은 다항식 뉴런(Polynomial Neuron)을 의미하며, 다항식 뉴럴 네트워크의 기본 연산 단위이다.다항식 뉴럴 네트워크는 기본 연산 단위인 다항식 뉴런을 층별로 쌓아 올려 구성된 네트워크이다. 다항식 뉴런의 구조는 표 1과 같은 다양한 회귀 다항식을 이용한다.

표 1 다항식 뉴런의 구조

Table 1 Structure of polynomial neuron

Type	Polynomial Neuron Structure
Linear	$y = a_0 + a_1 x_i + a_2 x_j$
Quadratic	$y = a_0 + a_1 x_i + a_2 x_i + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2$
Modified Quadratic	$y = a_0 + a_1 x_i + a_2 x_i + a_3 x_i x_j$

다항식 뉴럴 네트워크의 알고리즘 흐름은 아래와 같다.

[Step 1] 시스템 입력변수 결정

출력 변수 y에 관계하는 임의로 결정된 n개의 시스템 입력변수를 결정한다.

[step 2] PNN의 구조 결정

네트워크의 layer 수, 입력 변수의 수를 결정한다.

[step 3] 입력 변수와 다항식의 차수 결정

입력 변수의 수와 다항식의 차수를 기반으로 전체 입력 변수를 이용하여 생성할 수 있는 모든 조합 (S) 을 이용하여 현재 층의 Polynomial Neuron (PN)의 구조를 결정한다.

$$S = pCq \cdot T = \frac{p!}{(p-q)!q!} \cdot T \tag{1}$$

여기서, S는 중복되지 않는 PN의 경우 수를 의미하고, p는 현재 층의 입력으로 사용되는 전 층의 출력변수의 수 (단, 현재 층이 1층이면 주어진 데이터의 입력 변수의 수)를 나타내고, q는 step 2에서 결정된 기본 노드 PN에 입력으로 사용될 입력 변수의 수를 의미하며, T는 다항식의 종류 수를 나타낸다.

[step 4] 다항식의 파라미터 결정을 위한 LDA적용

각 뉴런의 다항식의 계수를 추정하기 위하여 (2)를 이용하여 각 다항식의 계수 결정한다.

$$w = S_w^{-1}(m_1 - m_2) \tag{2}$$

여기서, $S_i = \sum_{\mathbf{x} \in D_i} (\mathbf{x} - \boldsymbol{m_i})(\mathbf{x} - \boldsymbol{m_i})^T$, $S_W = S_1 + S_2$,

$m_i = 1/n_i \sum_{\mathbf{x} \in D_i} \mathbf{x}$이다.

[step 5] 성능 평가 및 노드의 선택

각 노드의 평가를 위해 패턴 분류 성능이 아닌 클래스 분리 기준인 LDA의 목적 함수 (3)를 사용하여 뉴런의 성능을 평가한다. 평가된 모든 노드들의 평가 값을 이용하여 임의로 결정된 개수 만큼 선택하여 다음 층의 입력 변수로 사용한다.

$$J = \frac{w^T S_B w}{w^T S_w w} \tag{3}$$

[step 6] 종료 판정

[step 5]에서 선택된 노드를 새로운 입력으로 하고 [step 3]~[step 5] 과정을 반복하며 새로운 층을 형성한다.

위 과정을 반복하여 생성된 네트워크의 층수가 설계자에 의해 미리 지정된 층수와 같으면 알고리즘을 종료한다

3. Support Vector Machine 기반 다항식 뉴럴 네트워크

다항식 뉴럴 네트워크의 과적합 문제를 해결하여 네트워크의 일반화 성능을 개선하기 위하여, SVM을 다항식 뉴럴 네트워크의 기본 구조로 사용한다. SVM을 다항식 뉴럴 네트워크의 기본 구조로 사용함으로써 다항식 뉴럴 네트워크의 일반화 성능 향상을 추구한다.

3.1 Support Vector Machine

SVM은 커널 매핑 및 최적화 기법을 이용하여 통계적 학습방법과 융합한 패턴 분류 알고리즘으로 다른 패턴 분류 알고리즘에 비해 비교적 간단한 구조를 가지며, 우수한 예측 성능을 보이는 것으로 알려져 있다. SVM은 특징 공간 (Feature Space)에서 초평면(Hyper-Plane)과 가장 인접한 지지벡터(Support Vector)들과의 거리인 마진(Margin)을 최대로 하는 최대 마진 초평면(Maximum Margin hyper-Plane)을 찾아서 이진 분류를 수행한다[7-11]. 다시 말하면, SVM 패턴 분류기는 입력공간상에서 패턴 분류를 위한 초평면(hyper-plane) $f(x)$의 coefficients w, b를 최적화 한다. 일반적으로 SVM에서 사용되는 초평면은 다음과 같다.

$$w \cdot x + b = 0. \tag{4}$$

여기서, $b \in R$는 모든 실수의 집합에 속한 임의의 값을 의미하고, $w \cdot x$는 계수 $w \in R^n$와 입력 데이터 x와의 벡터 내적값을 의미한다.

식 (4)와 같이 정의된 초평면을 이용하여 입력 데이터의 해당 클래스 레이블을 결정하기 위하여 아래 식을 이용한다.

$$\hat{y} = \begin{cases} 1, & w \cdot x + b \geq 1 \\ -1 & w \cdot x + b \leq 1 \end{cases} \tag{5}$$

여기서, $\hat{y}$는 입력데이터 x의 추정된 클래스 레이블을 의미한다.

그리고 식 (4)와 같이 표현된 초평면을 클래스 경계면(Decision Boundary Surface)라 부른다.

각 클래스에 속한 데이터들 중에서 초평면과 가장 가까운 거리에 있는 데이터 쌍의 거리를 최대 마진 (Maximum Margin)이라 하고, 아래식과 같이 표현한다.

$$Margin = \frac{2}{\| w \|} \tag{6}$$

만약 두 개의 클래스 분류 문제에서 주어진 데이터들이 선형 초평면으로 완벽하게 분리되어지지 않을 경우, 식 (5)를 (7)과 같이 변형하여 사용한다.

$$\hat{y}_j = \begin{cases} 1, & w \cdot x_j + b \geq 1 - \xi_j \\ -1 & w \cdot x_j + b \leq 1 - \xi_j \end{cases} \tag{7}$$

여기서, ξ_j는 j번째 변수에 해당되는 슬랙변수를 나타내며, 항상 0보다 크거나 같아야 한다($\xi_j \geq 0$).

이를 위한 일반적인 SVM의 목적함수는 (8)과 같다.

$$\min_{W,\xi} J_{SVM} + C\sum_{i=1}^{N} = \min_{W,\xi} \frac{1}{2} \| W \|^2 + C\sum_{i=1}^{N} \xi_i \tag{8}$$

$$s.t. \quad y_i(W^T X_i + b) \geq 1 - \xi_i, \forall i$$
$$\xi_i \geq 0, \forall i$$

여기서, C는 오분류와 패턴 분류 성능간의 trade-off를 나타내는 파라미터로 “trade-off 파라미터”로 정의한다.

일반적으로 선형식을 기본으로 하는 Linear SVM으로는 실제 현장에 적용하여 비선형 분류 문제를 해결하기에는 부족함이 있어서 Linear- SVM을 Nonlinear-SVM을 확장하여 사용할 필요가 있다 [12]. 이를 위해 표 2에 나열된 것과 같은 Kernel

표 2 Kernel function의 종류
Table 2 Types of kernel functions

Kernel Type	Kernel Function
Linear	$K(x,y)=x\cdot y$
Polynomial	$K(x,y)=(x\cdot y)^p$
Radial Basis Function	$K(x,y)=e^{-\parallel x-y\parallel^2/2\sigma^2}$

function $K(x_i,x_j)$를 추가하여 사용한다.

3.2 Support Vector Machine 기반 다항식 뉴럴 네트워크

SVM 기반 다항식 뉴럴 네트워크는 네트워크의 기본구조인 뉴런으로 SVM을 적용하는 네트워크이며, 뉴럴 네트워크의 동적생성은 앞에서 설명한 일반적인 다항식 뉴럴 네트워크 생성 방법을 변경없이 이용한다.

일반적인 SVM의 경우 주어진 입력 변수 전체를 입력으로 사용하지만, 다항식 뉴럴 네트워크의 기본 노드로 사용되는 SVM의 경우에는 전체 입력 변수들 중에서 사용자가 미리 정한 개수만큼의 입력 변수만을 입력으로 선택하여 사용한다. 그리하여, 전체 입력 변수의 개수가 p이고 q개의 입력 변수를 선택하여 입력 변수로 사용할 경우, 만들어질 노드의 경우의 수는 (9)와 같다.

$$S=pCq\cdot K=\frac{p!}{(p-q)!q!}\cdot K \tag{9}$$

여기서, K는 표 2에 보인 커널 함수의 종류의 개수를 나타낸다.

노드를 선택하기 위한 기준으로 학습데이터에 대한 오분류율과 SVM의 목적함수 식 (8)을 결합하여 (10)과 같은 선택함수를 정의하여 사용한다.

$$Criterion=\alpha\cdot MCR+(1-\alpha)\cdot J_{SVM} \tag{10}$$

여기서, MCR은 오분류율 (Mis-Classification Rate)을 의미하며, J_{SVM}은 식 (8)에 보인 $J_{SVM}=\frac{1}{2}\parallel w\parallel^2$을 의미한다.

SVM 기반 다항식 뉴럴 네트워크의 구조는 그림 2와 같다.

그림 2에 보인 것처럼 $(l-1)$층에서 선택된 노드의 출력이 l층의 입력 변수로 사용되어진다. 그림 3은 SVM 기반 다항식 뉴럴 네트워크의 구조 생성과정을 보인다. 각 레이어의 구조를 결정하기 위하여 각 노드의 패턴 분류 성능을 평가하고, 패턴 분류 성능을 기반으로 선택 과정을 거쳐 미리 정해진 개수만큼의 노드들이 선택된다.

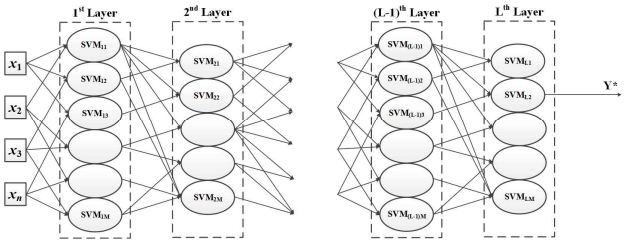


그림 2 SVM 기반 다항식 뉴럴 네트워크의 구조
Fig. 2 Architecture of SVM based polynomial neural networks

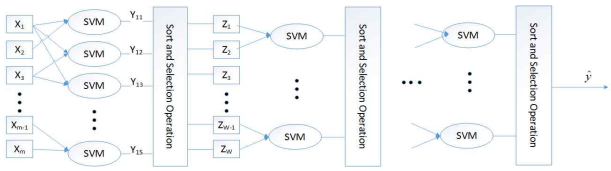


그림 3 SVM 기반 다항식 뉴럴 네트워크의 동적 생성
Fig. 3 Dynamic generation of SVM based polynomial neural networks

4. 다중 목적 입자 군집 최적화 알고리즘을 이용한 네트워크 구조 최적화

본 논문에서는 제안된 SVM 기반 다항식 뉴럴네트워크의 구조를 최적화하기 위하여 네트워크의 각 층의 노드 선택을 위하여 다중 목적 군집 최적화 알고리즘을 적용한다. 다중 목적 입자 군집 최적화 알고리즘의 목적함수는 네트워크의 기본 노드인 SVM의 기본 목적함수인 $J_{SVM}=1/2\parallel W\parallel^2$와 패턴 분류기의 학습 데이터에 대한 오분류율로 정의하여 사용한다.

4.1 다중 목적 입자 군집 최적화 알고리즘(Multi-Objective Particle Swarm Optimization: MOPSO)

MOPSO의 기본 알고리즘인 입자 군집 최적화 알고리즘(Particle Swarm Optimization: PSO)은 물고기 혹은 새와 같은 무리를 짓는 동물들이 먹이를 찾아 생존하는 행동양식을 모사하여 최적화 문제를 해를 탐색하는 방법이다. Genetic algorithm, Genetic Programming과 같은 알고리즘에 비해 PSO에서 사용되는 연산자들이 매우 간단한 수학적 연산자이기 때문에 최적화 알고리즘의 계산속도 측면에서 매우 유리한 조건을 가지고 있다. 또한 확률적인 방법에 기반을 두고 최적화 문제를 해결하는 다른 알고리즘에 비해 안정적이고 빠른 수렴 특징을 가진다고 알려져 있다 [13].

단일 목적 PSO의 장점을 보유 하면서 다중 목적 함수를 처리할 수 있는 MOPSO가 2002년 Coello에 의해 개발 되었다 [14]. 다중 목적함수 최적화 문제(Multi-objective Optimization Problem: MOP)는 아래와 같이 정의된다.

$$\begin{aligned}&\text{Optimize } f_i(\mathbf{X}),\ \text{for } i=1,\cdots,M\\&\text{subject to } g_j(\mathbf{X})>0,\ \text{for } j=1,\cdots,N\\&\qquad h_k(\mathbf{X})=0,\ \text{for } k=1,\cdots,K\\&\text{where, } \mathbf{X}=(x_1,\cdots,x_n),\ x_l^{(L)}\leqq x_l\leqq x_l^{(U)},\ l=1,\cdots,n.\end{aligned}$$

여기에서 f_i는 개별 목적함수(objective function). g_j와 h_k는 각각 부등식과 등식 제한조건(constraint)이며, M은 목적함수의 개수를 의미하고, N은 부등식 제한조건의 개수, K는 등식 제한조건의 개수를 나타낸다. 또한 $x^{(L)}$과 $x^{(U)}$는 입력 변수 x의 최솟값과 최댓값을 나타낸다.

다중 목적 함수 최적화 문제에서 해공간에서 하나의 해 $\mathbf{X}=(x_1, \cdots, x_n)$는 이에 상응하는 목적함수 공간상의 점 $f(\mathbf{X})=[f_1(\mathbf{X}), \cdots, f_M(\mathbf{X})]$을 갖는다. 해공간에서 서로 다른 두 점의 우열을 비교하여 우수한 해를 결정한다. 일반적으로 두 점 사이의 우열관계를 비교할 경우 "Dominance"관계를 이용한다.

해공간에서 다른 두 점 사이의 우열 관계를 비교하기 위한 "Dominance" 관계는 다음과 같다.

> 목적함수 $f_i(\mathbf{X})$를 최소화 하고자 할 경우,
> 해공간의 서로 다른 점 $\mathbf{X}_1$과 $\mathbf{X}_2$의 경우 $\mathbf{X}_1$이 $\mathbf{X}_2$보다 우수한 해라고 평가되기 위해서는 아래 두 조건을 모두 만족하여야 한다.
>
> $$\forall i \in 1, \cdots, M,\ f_i(\mathbf{X}_1) \leq f_i(\mathbf{X}_2), \qquad (11)$$
> $$\exists j \in 1, \cdots, M,\ f_j(\mathbf{X}_1) < f_j(\mathbf{X}_2)$$

위와 같은 두 조건을 모두 만족 시킬 때, $\mathbf{X}_1$은 $\mathbf{X}_2$를 "dominate"한다라고 할 수 있다.

만약 해공간상의 두 점 $\mathbf{X}_1$과 $\mathbf{X}_2$가 서로 dominate하지 않을 경우 두 점은 서로 무엇이 우수하다 할 수 없으므로 최적해들의 집합의 원소가 될 수 있다. 이와 같이 해공간상에서 dominate되지 않은 점들의 집합을 pareto 집합이라고 하며, 이를 pareto optimal 이라 부른다.

Pareto 집합을 이용한 MOPSO의 알고리즘은 다음과 같다.

step 1) 미리 정해진 수 만큼의 개체(particle: POP)들을 랜덤하게 생성한다.

(a) FOR i=0 TO MAX /* MAX: predetermined number of particles*/

(b) Initialize POP[*i*] randomly

step 2) 각 개체의 갱신에 필요한 속도를 초기화 한다.

(a) FOR i=0 TO MAX

(b) VEL[i]=0

step 3) 각 개체를 목적함수에 적용하여 적합도 (fitness)를 계산한다.

step 4) 평가된 각 개체의 적합도 벡터를 이용하여 dominated 되지 않은 개체들을 저장소 REP에 저장한다.

step 5) 현재까지 탐색한 탐색공간에 hypercube를 생성하고 각 개체들을 생성된 hypercube를 좌표 시스템으로 이용하여 위치시킨다. 여기서, 각 개체의 좌표는 각각의 목적함수의 값에 따라 결정된다.

step 6) 각 개체가 가지고 있는 메모리를 초기화 한다. 각 개체들이 보유하고 있는 메모리는 탐색공간에서 각 개체들이 탐색하기 위한 가이드 정보로 사용되어진다.

(a) FOR j=0 TO MAX

(b) *PBESTS*[j]=*POP*[j]

step 7) 식 (12)을 이용하여 각 개체의 이동 속도를 결정한다.

$$VEL[i] = W \cdot VEL[i] + R_1 \cdot (PBESTS[i] - POP[i]) + R_2 \cdot (REP[h] - POP[i]) \qquad (12)$$

여기서, W는 하중계수로 $0.4 \sim 0.9$ 사이의 값을 가지며 (본 논문에서는 $W=0.4$를 사용하였다.) R_1과 R_2는 가속상수로 $[0,1]$ 사이의 값으로 랜덤하게 결정된다. $PBESTS[j]$는 j번째 개체의 가장 우수한 위치를 나타낸다.

$REP[h]$는 저장소에 있는 개체들 중에서 얻어지는 값을 나타내며, h는 아래와 같이 결정한다.

하나 이상의 개체를 포함하고 있는 hypercube에 미리 정하여진 값 $(x>1)$을 해당 hypercube에 속한 개체들의 수로 나누어서 얻어진 값을 적합도로 할당한다(본 논문에서 $x=10$을 사용함). 이와 같이 각 hypercube에 할당된 적합도를 이용하여 roulette-wheel 선택 방법으로 hypercube를 선택한다. 선택된 hypercube에 속한 개체들 중에서 랜덤하게 하나의 개체를 선택하여 $REP[h]$라 한다.

step 8) 각 개체의 위치를 식 (13)를 이용하여 갱신한다.

$$POP[i] = POP[i] + VEL[i] \qquad (13)$$

step 9) 갱신된 개체들이 초기 탐색 범위($P_{\min}[i] \leq POP[i] \leq P_{\max}[i]$) 안에 존재하는지 확인하다. 만약 초기 탐색 범위 밖에 존재할 경우, 즉 $POP[i] < P_{\min}[i]$일 경우에는 $POP[i] = P_{\min}[i]$와 같이 설정하고, $POP[i] > P_{\max}[i]$일 경우에는 $POP[i] = P_{\max}[i]$로 설정한 후 이동 속도에 -1을 곱하여 식 (13)를 반복한다.

step 10) 각 개체의 위치 정보를 이용하여 목적함수의 값을 구한다.

step 11) 새롭게 얻어진 각 개체의 목적함수 값들을 이용하여 저장소 *REP*에 존재하는 개체들의 목적함수 값들과 비교하여 *REP*를 갱신한다.

step 12) 각 개체의 과거 $PBESTS[i]$와 현재 $POP[i]$의 적합도를 비교하여 현재 $POP[i]$의 적합도가 우수하면 $PBESTS[i]$를 아래와 같이 갱신한다.

$$PBESTS[i] = POP[i] \qquad (14)$$

$PBESTS[i]$와 $POP[i]$의 적합도를 비교할 때 앞에서 설명한 Pareto Dominance를 기준으로 판단하게 된다.

step 13) 종료조건을 검사하여, 종료조건을 만족하면 최적화 알고리즘을 종료하고, 만족하지 않았다면 step 7)로 이동하여 알고리즘 과정을 반복한다.

4.2 다중 목적 입자 군집 최적화 알고리즘을 이용한 SVM 기반 다항식 뉴럴 네트워크 최적화

제안된 SVM 기반 다항식 뉴럴 네트워크의 각 층의 구조를 최적화하기 위하여 앞에서 설명한 다중 목적 입자 군집 최적화 알고리즘을 이용한다. MOPSO를 이용하여 SVM기반 다항식 뉴럴 네트워크의 각 레이어 구조를 최적화하기 위한 Particle의 구조는 그림 4와 같다.

그림 4에서 보인 바와 같이 각 particle은 $(m+1)$차원의 해공간에서 탐색을 진행한다. 여기서 m은 입력 변수의 개수를 의미한다. MOPSO의 particle의 구조를 살펴보면, 앞 m개의 부분은 기본노드인 SVM의 입력 변수를 선택하기 위하여 사용되며, 마지막 하나는 SVM에 사용될 커널 함수의 종류를 결정하기 위하여 사용된다. 앞 m개의 부분은 $[0,1]$ 사이의 임의의 랜덤 값을 가지도록 초기화 되며, 마지막 부분은 $\{1,2,3\}$ 중에서 하나의 값을 임의로 선택한다.

MOPSO의 particle을 이용한 변수 선태과정은 그림 4와 같다. 그림 5에서는 5개의 입력 변수 $m=3$들 중에서 SVM 기반 다항식 뉴럴 네트워크의 기본 노드인 SVM의 입력으로 사용될 3개의 변수 $(l=1)$를 선택하고자 한다. 그림 4에서 빨간색으로 표시된 부분이 선택된 입력변수와 커널함수의 종류를 나타낸다.

MOPSO를 이용하여 제안된 SVM 기반 다항식 뉴럴 네트워크의 구조 최적화를 위한 과정은 그림 6과 같다. 그림 6에서 보인 바와 같이 SVM 기반 다항식 뉴럴 네트워크의 각 레이어의 구조는 MOPSO에 의해 결정된다. 단일 목적함수를 가지고 최적화 과정을 진행하는 Single-Objective PSO와 달리 MOPSO는 2개 이상의 목적함수의 값에 따라 우수한 particle을 평가하여 최적화를 진행한다.

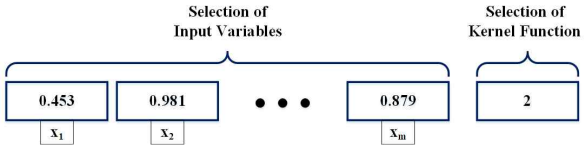


그림 4 SVM 기반 다항식 뉴럴 네트워크의 구조 최적화를 위한 Particle 구조

Fig. 4 Structure of particles for structural optimization of SVM based polynomial neural networks

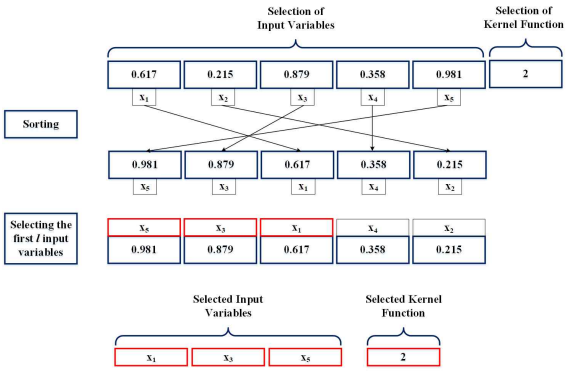


그림 5 Particle 정보해석

Fig. 5 Decoding of Particle Information

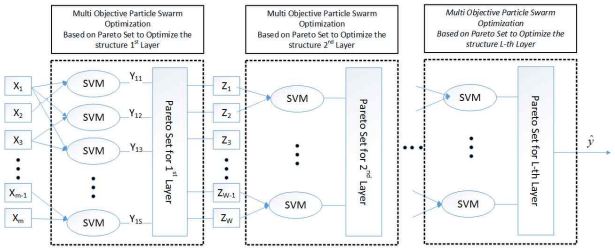


그림 6 제안된 패턴 분류기의 네트워크 구조의 최적화 과정

Fig. 6 Structure optimization procedure of proposed pattern classifier

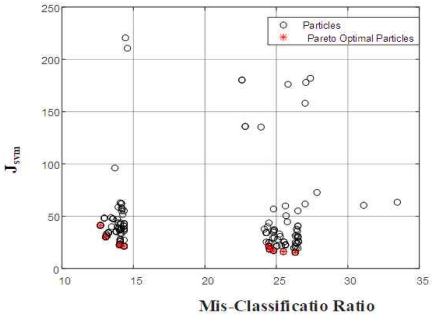


그림 7 MOPSO을 통해 획득한 Pareto set

Fig. 7 Pareto set obtained by MOPSO

최종적으로 모든 particle에 대해 2개 이상의 목적함수 측면에서 우수한 앞에서 설명한 pareto optimal set을 결정하게 된다. 이와 같은 과정을 거쳐 얻은 Pareto optimal set의 원소를 네트워크의 레이어를 구성하기 위한 노드로 사용하게 된다. MOPSO를 이용하여 제안된 SVM 기반 다항식 뉴럴 네트워크의 구조를 최적화하는 과정에서 Pareto Set을 정의하고 Repository에 저장한다. Repository에 저장된 Pareto set은 최적화를 위한 반복과정에서 갱신된다.

본 연구에서 사용된 2가지의 목적함수는 J_{SVM}과 오분류율이며 이 두 가지 목적함수를 이용하여 정의한 Pareto Set은 그림 7과 같다.

5. 실험 결과 및 고찰

본 연구에서는 다항식 뉴럴 네트워크의 기본 구조로 SVM을 적용한 SVM 기반 다항식 뉴럴 네트워크 패턴 분류기 설계 방법을 제안한다. 일반적인 SVM 기반 다항식 뉴럴 네트워크는 다항식 뉴럴 네트워크의 동적 생성 방법을 통해 네트워크의 구조를 레이어 별로 결정한다. 또한 최적화 알고리즘인 MOPSO를 적용할 경우, 제안된 패턴 분류기의 네트워크를 최적화 알고리즘을 이용하여 결정한다. 제안된 설계 방법을 통하여 얻어진 SVM 기

반 다항식 뉴럴 네트워크 패턴 분류기의 패턴 분류 성능을 평가하기 위하여 다수의 기계 학습 데이터를 사용하였다. 본 연구에서 수행된 실험은 통계적 유의성을 얻기 위하여 5 fold cross validation 방법에 따라 학습 데이터와 테스트 데이터로 나누어 수행되어진다.

표 3은 SVM 기반 다항식 뉴럴 네트워크의 패턴 분류 성능을 평가하기위하여 사용된 기계학습 데이터의 특성을 보인다. 본 논문에서 제안된 패턴 분류기의 성능을 평가하기 위하여 사용된 기계학습 데이터는 모두 2진 분류 데이터들이다.

표 3 기계 학습 데이터
Table 3 Machine learning data sets

Data	데이터 개수	입력변수의 수
Australian	690	14
Liver	345	6
German	1000	24
Heart	270	13
Diabetes	768	8

5.1 SVM 기반 다항식 뉴럴 네트워크

제안된 SVM 기반 다항식 뉴럴 네트워크는 그림 3에 보인바와 같이 각 노드의 패턴 분류 성능 평가와 패턴 분류 성능을 기반으로 하는 선택 과정을 거쳐 네트워크의 구조가 결정되어진다.

표 4는 5-fold cross validation을 이용하여 실험한 결과를 보인다. 패턴 분류 성능을 평가하기 위하여 패턴 분류기의 분류율을 평가 지수로 사용한다.

$$PCR = \frac{1}{N}\sum_{n=1}^{N} g(y_n, \hat{y_n}) \tag{15}$$

여기서, PCR은 Pattern Classification Ratio를 의미한다. 또한 $g(a,b)$는 식 (16)와 같다.

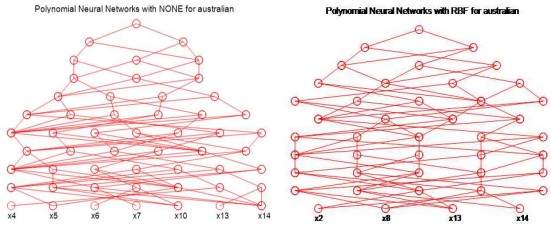
$$g(a,b) = \begin{cases} 1, & \text{if } a = b \\ 0, & \text{if } a \neq b \end{cases} \tag{16}$$

SVM 기반 다항식 뉴럴 네트워크를 이용한 머신러닝 데이터에 대한 패턴 분류 성능은 표 4에 보인다.

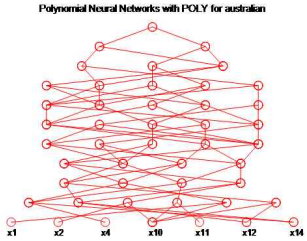
Australian 데이터의 경우, 학습과정을 통해 동적으로 생성된 SVM 기반 다항식 뉴럴 네트워크의 구조는 그림 8과 같다.

표 4 제안된 패턴 분류기 성능 비교
Table 4 Comparison of pattern classification performance

Data	Proposed Classifier	LVQ	SVM
Australian	79.71	68.9	65.0
Liver	69.28	66.4	68.1
German	74.0	28.7	70
Heart	84.07	66.0	60.8
Diabetes	76.69	74.0	73.7



(a) Linear Kernel Function (b) RBF kernel Function



(c) Polynomial Kernel Function

그림 8 동적 생성된 SVM 기반 다항식 뉴럴 네트워크 구조
Fig. 8 Structure of SVM polynomial neural networks generated dynamically for Australian data

5.2 MOPSO를 이용한 SVM 기반 다항식 뉴럴 네트워크 최적화

SVM기반 다항식 뉴럴 네트워크의 구조를 최적화하기 위하여 MOPSO 사용한다. 이를 위하여 사용된 다중 목적 입자 군집 최적화 알고리즘의 초기 설계 파라미터와 최적화 할 해공간의 탐색 범위를 표 5에 나타낸다.

표 5 다중 목적 입자 군집 최적화 알고리즘의 설계 파라미터와 탐색 범위
Table 5 Design parameter and search range of MOPSO

Parameters of MOPSO	
Maximum Generation	50
Swarm Size	50
Size of Repository	100
Acceleration Constant (c_1, c_2)	2.0
Random Constant (r_1, r_2)	[0, 1]
Inertia Weight (w)	[0.4 0.9]
Mutation rate	0.3
Search Space	
Input variables	$x_1, x_2, \cdots, x_m$
Kernel Function	1:(Linear Kernel), 2:(RBF kernel), 3:(Polynomial kernel)

학습 데이터의 패턴 분류 결과는 층을 증가시킬수록 개선이 되지만, 테스트 데이터의 경우 층의 개수가 증가 할수록 패턴 분류기의 복잡도가 증가되고, 패턴 분류기의 증가된 복잡도가 과적합 현상을 발생시켜 테스트 데이터에 대한 일반화 성능이 나빠짐을 알 수 있다. MOPSO에 의해 최적화된 SVM 기반 다항식 뉴

럴 트워크의 패턴 분류 성능과 Generic SVM 기반 다항식 뉴럴 네트워크의 성능을 다른 패턴 분류기와 비교한 결과를 표 6에 나열하였다.

표 6 제안된 패턴 분류기 성능 비교
Table 6 Comparison study of proposed classifier

Data	Proposed Classifier	Proposed Classifier optimized by MOPSO	KNN*[15]	PART*[15]
Australian	79.71	84.11	81.45	83.04
Liver	69.28	66.22	61.16	63.19
German	74.0	75.0	66.8	70.1
Heart	84.07	83.33	75.56	77.30
Diabetes	76.69	77.44	70.31	74.22

*: WEKA에 의해 얻어진 결과

제안된 패턴 분류기의 성능과 기존 연구된 패턴 분류기의 성능을 비교할 경우 기존 패턴 분류기에 비해 우수한 성능을 보임을 알 수 있다.

6. 결 론

본 논문에서는 동적 생성되는 특성을 가진 다항식 뉴럴 네트워크를 확장하여 네트워크의 기본 구조인 뉴런에 Support Vector Machine을 적용하여 SVM 기반 다항식 뉴럴 네트워크를 제안하고 이를 패턴 분류기로 적용하였다. SVM은 노이즈에 강인한 패턴 분류기로 알려져 있으며, 대표적인 2진 분류 문제에 대한 패턴 분류기로 인식되어진다. 이와 같은 SVM의 장점을 이용하여 네트워크를 구축하기 위하여 동적 생성 네트워크인 다항식 뉴럴 네트워크의 기본 노드로 SVM을 채택하여 패턴 분류기를 설계한다. 동적 생성 네트워크는 매 층마다 네트워크의 구조를 생성하는 과정을 거쳐 최종 네트워크의 구조를 결정한다.

이를 확장하여 다중 목적함수를 가진 다중 목적 입자 군집 최적화 기법을 사용하면, 두 가지의 선택기준을 하나로 통합하지 않고 두 가지를 동시에 고려하여 Pareto set을 형성하여 SVM 기반 다항식 뉴럴 네트워크의 각 층의 구조를 최적화 한다.

제안된 설계 방법에 의하여 설계되고 최적화된 SVM 기반 다항식 뉴럴 네트워크의 성능을 확인하기 위하여 다양한 기계학습 데이터에 이를 적용하여, 패턴 분류 성능을 얻었고, 이를 이미 연구되어진 다른 패턴 분류기와 비교 평가하였다. 패턴 분류 성능의 비교 평가를 통해 실험결과에 보인바와 같이 기존 기계학습 알고리즘에 비해 우수한 결과를 얻을 수 있었다.

감사의 글

본 연구는 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2017R1D1A1B03032333).

References

[1] G. B. Hwang, Q. U. Zhu, C. K. Siew, “Extreme Learning Machine: Thoery and Applicaitons”, *Neurocomputing*, vol. 70, pp. 489-501, 2006

[2] S.-K. Oh and Pedryza, W, “The Design of Self-Organizing Polynomial Neural Networks”, *Information Science*, vol. 141, pp. 237-258, 2002

[3] S.-K. Oh and W. Pedrycz, "Identification of Fuzzy Systems by means of an Auto-Tuning Algorithm and Its Application to Nonlinear Systems", *Fuzzy sets and Systems*, vol. 115, no. 2, pp. 205-230, 2000.

[4] S.-K. Oh, W. Pedrycz, and D.-W. Kim, "Hybrid Fuzzy Polynomial Neural Networks", *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 3, pp. 257-280, June, 2002.

[5] S.-K. Oh and W. Pedrycz, "The design of self-organizing Polynomial Neural Networks", *Information Science*, vol. 141, pp. 237-258, 2002.

[6] S.-K. Oh, W. Pedrycz and B.-J. Park, "Polynomial Neural Networks Architecture: Analysis and Design", *Computers and Electrical Engineering*, vol. 29, Issue 6, pp. 703-725, 2003.

[7] C. Cortes and V. N. Vapnik, “Support-vector networks”, *Mach. Learn.*, vol. 20, no. 3, pp. 273-297, 1995

[8] T. B. Trafalis and H. Ince, “Support vector machine for regression and applications to financial forecasting”, *in Proc. IEEE-INNSENNS Int. Joint Conf. Neural Netw.*, vol. 6, pp. 348-353, 2000

[9] W. S. Noble, “Support vector machine application in computational biology”, *in Kernel Methods in Computational Biology, B Schokopf, K. Tsuda, and J.-P. Vert, Eds. Cambridge*, MA, USA: MIT Press, 2004

[10] K.-S. Goh, E.-Y. Chang, and B.-T. Li, “Using one-class and two-class SVMs for multiclass image annotation”, *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 10, pp. 1333-1346, 2005

[11] D. Isa, L.-H. Lee, V. P. Kallimani, and R. RajKumar, “Text document preprocessing with the Bayes formula for classification using the support vector machine”, *IEEE Trans.Knowl. Data Eng.*, vol. 20, no. 9, pp. 1264-1272, 2008

[12] M. B. Karstan, “Kernel methods in bioinformatics”, in Handbook of Statistical Bioinformatics, *Part 3, New York, NY, USA: Springer*, pp. 317-334, 2011

[13] R. E. Perez and K. Behdinan, “Particle swarm approach for structural design optimization”, *Computer & Structures*, vol. 85, Issues 19-20, pp. 1579-1588, 2007

[14] C. A. Coello Coello, and M. S. Lechuga, "MOPSO: A

Proposal for Multiple Objective Particle Swarm Optimization," *Proceedings of the 2002 IEEE Congress on Evolutionary Computation*, pp. 12-17.

[15] E. Frank, M. A. Hall, and I. H. Witten, The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", *Morgan Kaufmann, Fourth Edition*, 2016.

저 자 소 개

노 석 범 (Seok-Beom Roh)

1997년: 원광대학교 컴퓨터공학과 석사
2006년: 원광대학교 제어계측공학과 박사
2016년~현재 : 수원대학교 전기공학과 연구교수. 〈관심분야〉 기계학습, 퍼지 시스템, 퍼지-뉴럴 네트워크 등.
Phone: +82-31-222-6544
E-mail: rsb@suwon.ac.kr

오 성 권 (Sung-Kwun Oh)

1981년 : 연세대학교 전기공학과 공학사. 1983년~1989년 : 금성산전연구소(선임연구원). 1993년 : 연세대학교 전기공학과 공학박사. 1996년~1997년 : 캐나다 Manitoba 대학 전기 및 컴퓨터 공학과 Post-Doc. 1993년~2004년 : 원광대학교 전기전자 및 정보공학부 교수. 2005년~현재 : 수원대학교 전기전자공학부 교수. 2002년~현재 : 대한전기학회 및 한국지능시스템학회 편집위원. 2013년~현재 : Information Sciences 편집위원. 〈관심분야〉 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 Computational Intelligence, 지능제어 등.
Phone : +82-31-229-8162
E-mail : ohsk@suwon.ac.kr