

論文96-33B-1-21

단일특징 분할 회귀트리의 학습성능 개선을 위한 회귀신경망

(Regression Neural Networks for Improving the Learning Performance of Single Feature Split Regression Trees)

林 淑*, 金 聖天**

(Sook Lim and Sung Chun Kim)

요 약

본 논문은 회귀트리에 기반을 둔 회귀 신경망을 제안한다. 회귀트리를 세개의 계층을 갖는 전향 신경망에 사상하고, 첫번째 계층에 다중특징 분할함수를 형성시켜 신경망이 보다 더 최적인 입력 공간의 분할을 갖도록 한다. 본 연구에서는 신경망 트레이닝을 위한 두 가지 지도 학습 알고리즘을 제안하여 단일특징 분할함수와 다중특징 분할함수에 실험한다. 실험결과, 제안된 회귀 신경망은 기존의 단일특징 분할 회귀트리 및 단일특징 분할 회귀신경망보다 학습능력이 우수함을 입증한다. 또한 본 논문에서 제안한 알고리즘이 학습 능력을 저하시키지 않으면서도 효과적으로 과성장한 회귀트리를 가지치기 할 수 있음을 보인다.

Abstract

In this paper, we propose regression neural networks based on regression trees. We map regression trees into three layered feedforward networks. We put multi feature split functions in the first layer so that the networks have a better chance to get optimal partitions of input space. We suggest two supervised learning algorithms for the network training and test both in single feature split and multifeature split functions. In experiments, the proposed regression neural networks is proved to have the better learning performance than those of the single feature split regression trees and the single feature split regression networks. Furthermore, we shows that the proposed learning schemes have an effect to prune an over-grown tree without degrading the learning performance.

I. 서 론

입력값을 유사성에 따라 분류하는 문제는 패턴인식에서의 분류문제(classification problem)와 회귀분석(regression analysis)에서의 회귀문제(regression

problem)문제로 대별된다^[1]. 분류문제에서의 작업이 관찰된 입력 데이터가 속하는 클래스(class)의 범주를 결정해 주는 것이라면, 회귀문제에서는 입력 데이터에 상응하는 출력변수의 값을 예측하는 것이다. 이들 두 문제에 접근하는 방법으로서 트리(tree)구조를 사용한 기법은 그 구조가 간단하고 명확하게 표현되는 점때문에 많은 연구가 진행되고 있다.

이러한 트리구조에 기초한 접근방법은 생성되는 트리크기에 따라 계산비용에 영향을 주므로 트리설계시, 주어진 작업을 수행하기에 최적인 크기를 찾는 것이 주요한 관심사의 하나이다. 트리 접근방법은 주어진 데이터에 대하여 한번 임계치(threshold value)가 결정

* 正會員, 麗水水産大學校 컴퓨터 工學科

(Dept. of Computer Engineering, Yosu National Fisheries Univ.)

** 正會員, 西江大學校 電子計算學科

(Dept. of Computer Science, Sogang Univ.)

接受日字: 1995年11月23日, 수정완료일: 1995年12月28日

되어 트리가 구축된 이후에는 새로운 데이터가 들어왔을 때 트리를 전면적으로 재구축해야 하는 점이 있다. 또한 잡음에 민감한 데이터에 대하여 일단 잘못된 경로가 선택되었을 때 오류를 회복하기가 어렵다는 한계가 있다.

이러한 결점에 대하여 계층구조를 갖는 신경망(layered neural networks)은 주어진 관찰 데이터에 대하여 네트워크에서 연속적으로 연결강도와 임계치를 수정해 주는 적응성(adaptability)이 있다. 반면 주어진 작업을 수행하기에 적당한 네트워크의 크기가 체계적으로 설계되기가 어렵다는 점때문에 신경망에서의 주요과제 또한 최적인 연결강도(connection strength)를 찾는 알고리즘을 개발하는 것뿐만이 아니라 층(layer)의 크기 및 각 층당 뉴런(neurons)의 수를 최적화시키는 것이 문제가 된다. 이러한 신경망 최적 구축을 위한 방법중에는 비대한(oversized) 신경망을 가지치는(pruning) 방식을 통하여 중복되는 뉴런과 그 연결을 줄여가는 연구가 있고 반면, 작은 신경망에서 출발한 후 점차 한 뉴런씩 확장해 가는 알고리즘, 그리고 트리구조를 신경망위에 재편성한 기법이 있다^[2, 3, 4, 5].

이들 중 주어진 문제의 크기에 보다 적합한 신경망 위상(topology)을 구축하기 위하여 트리구조를 사상(mapping)시키는 기법은 오류복구의 어려움이 있는 트리구조의 결점을 보완할 수 있고, 또한 빠른 학습시간을 갖는 트리구조의 장점을 살릴 수 있다는 점이 있다. 그러나 이와 관련한 기존의 연구에서는 트리구조를 그대로 사상시켰기 때문에 신경망위에서의 학습이 트리구조위에서의 학습 수행능력에 비하여 소폭 개선됨으로써 오히려 신경망으로의 사상에 대한 오버헤드(overhead)문제를 야기시킨다. 즉, 단일특징 분할 회귀트리(single feature split regression tree)에 있어서, 기존의 기법은 신경망위에서 연결강도를 교정한 이후에도 주어진 데이터를 분류하기 위한 경계선들이 항상 축에 평행함으로써 트리구조의 결점을 여전히 가지기 때문에, 신경망의 크기가 보다 적합한 형태로 조절되었음에도 불구하고 심화학습(further learning) 효과가 두드러지지 못하는 점이 있다.

따라서 본 논문은 트리구조를 그대로 사상시키는 것이 아니라 첫번째 은닉층(first hidden layer)에의 연결을 수정함으로써, 축에 항상 평행인 경계선만이 도출되지 않고 초다면체(hyperpolyhedron)의 형태인

경계선을 가지게 하여, 보다 최적인 경계선을 도출하고자 한다. 제안한 기법은 단일특징 분할 회귀트리를 다중특징 분할 회귀트리(multi feature split regression tree)로 확장시킨 효과를 갖는다. 학습하기 위한 알고리즘으로는 경쟁학습(competitive learning)을 적용하였는 바, 학습성능을 감소시키지 않으면서도 효과적으로 과성장(over grown)한 트리를 가지치기 할 수 있다. 두 가지 실험을 통하여 각 경우에 기존의 연구방법에 비해 제안한 회귀 신경망의 성능이 향상됨을 고찰한다.

본 논문은 다음과 같이 구성된다. 2장에서는 분류와 회귀문제에 적용되는 기존의 트리기법을 살펴보고, 3장에서는 제안한 네트워크를 구축하고 그 성능을 평가한 뒤, 4장에서는 두 가지 실험결과에 따른 결론을 요약한다.

II. 분류와 회귀문제에 적용되는 트리기법

관찰된 결과에 따라 임출력이 대응되어 있는 주어진 예들로부터, 입력공간과 출력공간을 각각 그 유사성에 따라 분류하여 최적의 군집(cluster) 개수를 찾는 문제를 고려해 보자. 패턴분류 문제에서, 일반적으로 출력은 그 패턴이 속하는 범주에 대한 분류표지(label)가 된다. 따라서 출력공간을 동질성에 따라 몇개의 그룹으로 나누고자 할 때, 다수를 차지하는 패턴이 한 그룹을 대표하게 되고, 엔트로피(entropy)와 같은 측도(measure)가 그 그룹의 동질성을 수치적으로 나타낸다. 회귀문제에서의 출력영역은 실수값으로 표현되는 부분집합으로서, 각 부분집합에 속해 있는 패턴들의 평균값이 한 그룹을 대표하고 분산과 같은 측도를 사용한다.

1. 트리 구조적 기법

이진 트리(binary tree)가 생성되는 방법과 같이 트리에 기초한 방법들은 전체집합을 두 개의 하행(descendant) 부분집합으로 분할시키고 다시 각 좌우 노드에 대해 각각 하행 부분집합으로 분할하는 것을 반복함으로써 구성된다. 분류문제에서 각 말단 노드(terminal node)들은 클래스 분류표지에 의해 명시된다. 두 개이상의 말단 노드들이 같은 클래스 분류표지를 가질 수 있다.

트리를 구축하는데 있어서의 문제는 얼마나 더 작은 조각들로 전체집합을 분할시킬 것인가를 결정하는 것

이다. 기본적인 개념은 하행 부분집합 각각에 속하는 여러 패턴들이 부모 노드(parent node)에서 섞여 있는 정도보다 덜 혼합되도록 분할을 선택하는 것이다.

트리를 구축하기 위한 작업은 다음의 요소로 구성된다.

- 1) 하행노드로 분할하기 위한 분할함수(split function)를 정의한다.
- 2) 말단노드의 형태를 결정한다.
 - a) 트리 구조적인 분류(tree structured classification)에서, 말단노드는 분류된 클래스에 대한 표시가 지정된다.
 - b) 트리 구조적인 회귀(tree structured regression)에서, 말단노드는 분류된 각 영역에 대한 대표값이 선정된다.
- 3) 현재 단계까지 생성된 말단 노드에 대하여 복잡도(impurity)가 최대로 감소되는 분할값을 선택한다.
- 4) 이 분할값에 대하여 데이터가 속하는 영역의 여부(yes, no)를 결정한다.
- 5) 한 노드가 말단노드로 선언되기 위한 규칙, 즉 한 노드가 단일 특징만으로 구성되면 종료시키거나 또는 다수인 특징(feature)이 있다하더라도 미리 지정된 오류허용 조건이 만족되면 종료시킨다.

결정트리(decision tree)는 위의 구성요소 중 노드 분할을 위한 분할 결정함수로서 엔트로피 측도를 이용하여 입력 데이터를 초평면(hyperplane)으로 나누는 트리 구조적인 분류자(tree structured classifier)이다.

한편, 회귀트리(regression tree)는 분산과 같은 통계량을 그 분할 결정함수에 이용하였으며 말단 노드의 형태는 통계량에 기초하여 산출된 상수값에 따라 영역이 구별되는 트리 구조적인 회귀(tree structured regression) 방법이다.

이들 트리 구조적인 접근방법들은, 단일 프로세서(processor) 위에서의 구현을 가정할 경우, 생성된 트리 크기에 종속된 계산속도를 가진다는 한계에 큰 관계없이 계층신경망에 비해 보다 빠른 학습속도를 가진다는 장점이 있다. 그러나 그 학습능성에 있어서는 일단 한 노드가 결정되었을 때 오류를 복구하기 위한 유연성이 떨어짐으로써 계층 신경망에 비하여 상대적으로

로 저하된 학습능력을 가진다.

1) 회귀 트리

결정트리가 분류문제에서 주로 사용하는 비인수적인 분류(non-parametric classification) 방법이라면, 회귀모형을 위한 회귀트리는 결정트리와 거의 동형의 구조를 가지는 비인수적인 회귀 접근방법이다.

회귀 트리를 생성하는 방법은 다음과 같다. 주어진 관찰 벡터들을 가능한 모든 그룹으로 나누어 각 그룹에 대한 분산값이 최소가 되는 입력값을 찾아 그 값을 분할값으로 정하여 두개의 영역으로 분할하고 분할된 각 영역에 대하여도 위의 방법을 반복하여 적용하면서 계속적인 분할영역을 생성해 나간다. 즉 트리의 노드가 확장되는 매 단계마다 분할에 따라 생기는 최소제곱의 오차값(least squares error)을 최대로 감소시키는 그 임계치를 구한다. 이와 같은 분할 측도는 다음의 식으로 요약된다.

$$t_j = \frac{1}{N} \sum (y_0 - \bar{y}_0)^2 - \left(\frac{1}{N_1} \sum (y_1 - \bar{y}_1)^2 + \frac{1}{N_2} \sum (y_2 - \bar{y}_2)^2 \right)$$

단, t_j 는 j 번째 특징축(feature axis)에 대한 임계치이고 N , N_1 및 N_2 는 현재 노드에서의 데이터 개수, 왼쪽 하행노드의 데이터 개수 및 오른쪽 하행노드의 데이터 개수를 의미한다. y_0 , y_1 및 y_2 는 현재노드와 각 하행노드에 속한 데이터들의 평균값인 \bar{y}_0 , \bar{y}_1 및 \bar{y}_2 에 대하여 주어진 관찰벡터에 대한 출력 변수들을 의미한다. 트리 확장 알고리즘은 최소 제곱 오차의 감소 정도가 사전에 정의된 값에 도달할 때 종료된다. 단일 회귀트리에서의 내부 노드들은 단일 결정트리에서와 같이 분할된 영역을 표현하며, 관찰 벡터를 이루는 각 특징별로 분할 함수가 사용된다. 말단 노드들은 그 말단 노드에 속하는 트레이닝 데이터의 평균으로서 예측하고자 하는 변수에 대한 상수값을 표현한다. 그림 1은 이차원 문제(x_1, x_2)에 대한 회귀트리의 예를 보여 준 것이다.

회귀에 기초한 트리의 중요한 성질 중 하나는 유한개의 트레이닝 예들의 집합으로부터 입력-출력 사이의 대응을 일반화할 수 있다는 것이다. 그러나 회귀트리의 주요 결점은 트리를 구성하는 매 단계마다 결정된 임계값이 각 단계에 따른 입력공간을 양분할(bipartition)하기 위한 최적 경계선을 도출하지 못한다는 것이다. 즉 트리를 구성하는 매 단계마다 결정된 임계값이 각 축에 대해 항상 평행이 되는 초평면을 구성함으로써 무작위로 분포되어 있는 입력공간을 분할하기에

는 적당하지 못하다. 또한 회귀 트리는 그림 1의 b)에서와 같이 플랫폼(flat)형의 결과가 도출되며 유연하지 못한 회귀표면을 가진으로써 연속적인 출력영역을 원하는 응용분야에 적당하지 못하다.

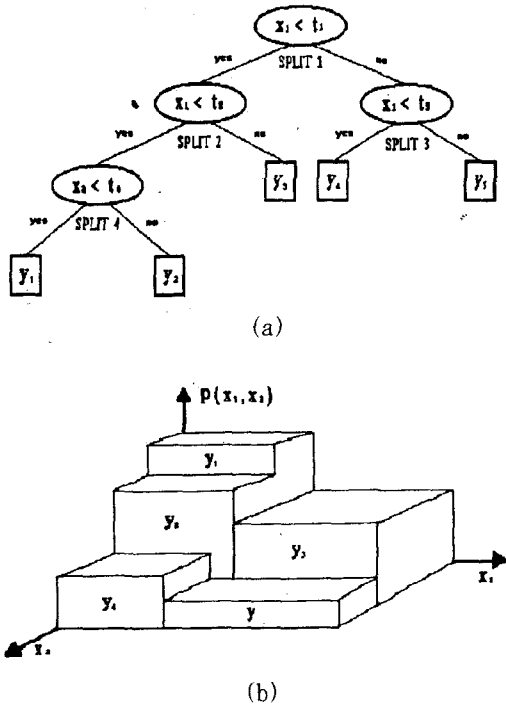


그림 1. a) 이차원(x_1, x_2)문제에 대한 회귀트리에 b) a)에 대한 회귀표면 단, $P(x_1, x_2)$ 는 근사 추정자, 예측 출력 변수 y_i 는 상수값을 가진다.

Fig. 1. a) Example of regression tree for a two dimensional problem on (x_1, x_2) . b) The regression surface corresponding to a), where $P(x_1, x_2)$ is a approximator, and predicted output variable y_i is a constant.

2) 신경망과 트리구조적인 기법의 결합

일반적으로 함수 추정을 위한 신경망 모델로는 전향 신경망(feedforward neural network), RBF(Radial basis function) 네트워크 및 CMAC(cerebellar Model Arithmetic Computer) 등이 있다. 전향 신경망에서는 역전파(backpropagation) 학습알고리즘을 일반적으로 사용하며, RBF 네트워크 및 CMAC 네트워크는 k-mean 군집알고리즘을 이용하여 각 군집의 중심값과 크기를 결정하고 각 군집당 출력 추정치는 LMS(least mean square) 규칙을 이용하여 학습한다^{16, 7, 8, 9, 10, 11}

회귀트리는 그 자체로 함수 추정을 위한 한 방법으로 여러 연구자들이 위에서 열거한 함수 추정을 위한 신경망 모델들과 퍼지 추론 시스템과의 결합을 시도하였다. Sethi와 Yu의 연구는 회귀트리를 엔트로피 네트워크에 적용시켰는데 전환된 네트워크의 세번째층에 많은 뉴런을 추가하여 출력값 추정 능력을 향상시켰다¹². Prager는 CART(classification and regression trees)와 CMAC를 결합하였는데 CART 회귀트리의 말단 노드들이 형성한 초상자(hyperbox)들을 각각 일정 비율로 확장시켜서 초상자 서로간에 겹치는 부분들을 형성함으로써 주어진 입력값에 대하여 다수의 말단노드들이 활성화하는 효과를 얻었다^{18, 13}. Lim과 Kim은 회귀트리에 기반을 둔 퍼지 추론 시스템(fuzzy inference system)을 제안하였는 바, 각 말단 노드에 속하는 트레이닝 데이터에 대한 입력 공간상의 평균과 분산 벡터등을 이용하여 각 회귀트리의 말단 노드들을 퍼지 연상 메모리(fuzzy associative memory)에 사상시켰다¹⁴.

계층신경망에서의 중요한 과제중의 하나는 주어진 문제를 해결하는데 적합한 적당한 크기의 네트워크가 체계적으로 결정되어 있지 못하다는 데 있다. 네트워크 크기가 결정된다는 것은 불필요한 노드와 연결수를 줄여줌으로써 효과적으로 입출력간의 대응이 일어나고 따라서 학습시간을 단축시킬 수 있음을 의미한다. 이러한 네트워크 구축문제는 작거나 혹은 큰 네트워크에서 출발하여 점차로 적당한 크기의 네트워크를 찾아가는 해결방안과, 트리에 기초한 분류나 회귀기법을 네트워크에 사상시키는 해결방안이 연구되었다.

특히 엔트로피 네트워크(entropy network)는 결정 트리를 3개층으로 이루어진 전향신경망에 사상하여 네트워크의 노드와 계층의 수를 체계적으로 결정한 방법이다¹². 사상된 네트워크는 기존의 각 층당 완전연결(full connection)로 이루어진 계층신경망에 비교했을 때 더 적은 수의 연결링크와 뉴런을 가진으로써 과도한 크기의 네트워크를 조정하는 효과를 가진다.

III. 다중특징분할(multi feature split) 효과를 위한 회귀 신경망

1. MRNN(Multi feature split Regression Neural Networks)
회귀트리를 구성할 때 한 임계치는 이웃한 두 분할

공간들의 가장 가까운 두 점의 중점이므로 따라서 표본 데이터가 전 공간에 걸쳐 분포되어 있는 경우를 제외하고는 각 분할에 속한 데이터의 분포에 대한 최적의 정보를 그 자체로 표현하기에는 부족하다. 이러한 사실은 그림 2를 통해서 관찰될 수 있다. 여기서 m_1, m_2 는 각 영역에 대한 출력벡터들의 평균값이고, n_1, n_2 는 이웃한 두 분할영역의 경계선에 대하여 가장 근접한 점이다.

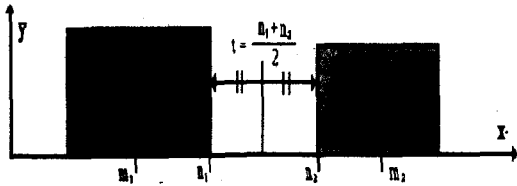


그림 2. 임계치(t)의 특성
Fig. 2. Property of threshold value t.

엔트로피 네트워크가 구축된 규칙과 유사하게, 회귀트리를 계층신경망에 사상시킴으로써 주어진 문제에 보다 적합한 네트워크 크기를 결정할 수 있다. 그러나 한 특징별로 분할함수를 적용시키므로 네트워크의 매 단계에서 고려되는 입력변수는 임계치와 해당 특징축의 변수만을 고려하게 된다. 따라서 네트워크위에서 학습한 효과는 한 축에 대하여 미세조정을 함으로써 크기가 감소되거나 커지는 초상자를 여전히 가지게 된다. 이것은 네트워크위에서 학습하여도 분할이 이루어질 때마다 항상 축에 평행한 경계선으로 조정되므로 무작위로 분포가 가해진 입력공간을 정확한 경계선으로 나누는데 한계가 있다. 따라서 심화학습을 한 후에도 회귀트리가 가지는 단점인 초상자의 형태를 그대로 가진다.

그림 3은 그림 1에서 보여준 회귀트리의 예를 본 연구에서 제안한 네트워크로 구성한 것이다. 이 네트워크를 MRNN(multi feature split regression neural networks)으로 지칭한다.

MRNN의 설계과정은 다음으로 요약된다.

- 1) 주어진 입력 벡터에 대하여 회귀트리를 생성한다. 회귀트리를 생성하는 방법은 2.2절에서 언급한 바와 같다.
- 2) 회귀트리의 내부 노드(internal node)들을 MRNN의 첫번째 은닉층의 뉴런으로 사상한다.

- 3) 회귀트리의 말단 노드(terminal node)는 MRNN의 두번째 은닉층의 뉴런으로 사상한다.
- 4) 첫번째 은닉층에 주어진 입력 벡터 전체를 연결한다. 따라서 회귀트리의 결과, 연결이 없었던 입력 벡터들과 첫번째 은닉층의 뉴런이 연결된다. 첫번째 은닉층에 대한 연결의 초기 가중치는 회귀트리의 결과를 이용한다.
- 5) 첫번째 은닉층과 두번째 은닉층사이의 연결방법은 회귀트리에서의 말단노드에 대해 경로가 있는 모든 내부노드들을 연결한다. 말단노드가 연결된 내부노드의 왼쪽 부분트리에 존재하면 -1, 오른쪽 부분트리에 존재하면 +1로 가중치를 정한다.
- 6) 세번째 은닉층에 대한 연결의 가중치는 두번째 은닉층의 활성화값으로 정하고, 두번째 은닉층의 뉴런에 한 개의 뉴런으로 구성된 출력층을 연결한다.

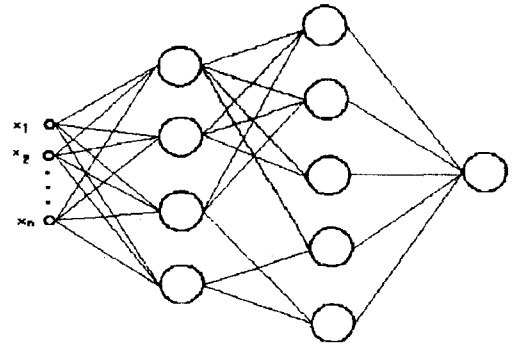


그림 3. MRNN. 첫번째와 두번째 은닉층의 뉴런은 각각, 그림 1에서 본 회귀트리의 내부노드와 말단노드에 해당된다. 세번째 층은 예측 출력변수를 위해 추가된 출력층이다.

Fig. 3. MRNN. First hidden layer neurons correspond to the internal nodes and second hidden layer corresponds to the terminal nodes of regression trees, in Fig. 1. Third layer is attached to second hidden layer for predicted output variable.

따라서 이 마지막 층은 출력변수의 함수적인 연속성을 유지시켜 회귀함수 추론기능을 향상시킨다. MRNN은 이와같은 세개 층으로 구성된 회귀모델을 위한 계층신경망이다.

2. 네트워크 트레이닝

주어진 트레이닝 데이터를 가지고 구축된 회귀트리로부터 네트워크의 초기 연결강도를 설정하고 구성한

다. 다중특징 분할 회귀트리의 반투명성을 유지하기 위해 네트워크의 첫째 은닉층을 학습한다. MRNN은 네트워크의 두번째 층에서 역전과 알고리즘에 의해 훈련된다. 두 번째 층에 대한 오차는 다음과 같이 정한다.

error = 1.0 - activation for winner neuron
 error = -1.0 - activation for looser neuron

단, 획득뉴런(winner neuron)은 가장 큰 활성화(activation) 값을 갖는다.

본 논문에서 제안한 기법은 획득뉴런을 정하는 다음의 두 가지 다른 실험방법을 통해 제안된 MRNN위에서 학습한 후 각각의 성능을 비교, 평가한다.

[실험 1]

주어진 데이터에 대한 회귀트리에 결과에 대해 별도의 조정없이 회귀결과에 의존적인 지도학습(supervised learning)방법을 사용한다. 즉, 회귀트리의 학습결과에 따라, 주어진 입력 데이터가 속한 말단노드에 상응하는 뉴런을 획득뉴런이 되도록 한다. ■

[실험 2]

회귀트리가 산출한 출력값들에 기초하여 주어진 학습 데이터의 출력 공간값에 가장 근접한 뉴런을 획득뉴런으로 정하기 위해 다음과 같은 Gaussian basis 함수를 정의한다.

$$g_c(y_p) = \exp \left[-\frac{(y_p - \bar{y}_c)^2}{2\sigma_c^2} \right]$$

단, y_p 는 주어진 트레이닝 데이터의 출력 변수이고, \bar{y}_c 및 σ_c^2 은 두번째 층에 있는 각 뉴런에 대한 출력값들의 평균 및 분산을 의미한다. 따라서 이 Gaussian basis 함수값이 최대치가 되는 뉴런이 획득뉴런으로 지정된다. ■

획득뉴런에 대한 출력값의 평균과 분산은 다음과 같이 조정한다.

$$\sigma_{wn}^{(new)} = \sigma_{wn}^{(old)} (1 - \alpha) + (y_p - \bar{y}_{wn}^{(old)})^2 \alpha$$

$$\bar{y}_{wn}^{(new)} = \bar{y}_{wn}^{(old)} (1 - \alpha) + y_p \alpha$$

$$\alpha = \rho / ((N/2^k) + 1)$$

단, ρ 는 학습계수(learning coefficient), N은 총 학습데이터의 개수를, k는 그 획득뉴런으로 들어 오는 연결 링크(connection links)의 수를 의미한다.

MRNN에서 최종층의 기대출력은 실험 1, 2에 대해 공통적으로 다음과 같은 가중평균으로 계산한다.

$$\hat{y} = \frac{\sum_c w_c \bar{y}_c}{\sum_c w_c}$$

단, \bar{y}_c 는 두번째 층의 각 뉴런에 속한 데이터들에 대해 상응하는 출력 데이터의 평균값이고, 가중치는 0에서 1사이의 값으로 정규화(normalization)된 두 번째층의 뉴런의 활성화값을 의미한다.

3. 성능 평가

MRNN을 포함한 각 회귀분석모델의 성능은 모든 입력 데이터에 대하여 출력 y_i 와 각 회귀모델의 추정치인 y_{ie} 에 대한 MSE(mean squared errors)로 측정한다.

$$MSE = (1/n \sum (y_{ie} - y_i)^2)^{1/2}$$

단, n 은 표본 데이터의 개수를 의미한다.

주어진 학습 데이터 집합은 4 개의 입력특징과 1개의 출력특징으로 구성되며 400개의 트레이닝 집합과 132개의 테스트 집합을 실험하였다. RT는 단일특징 분할회귀트리에서의 실험결과를 의미하며 SRNN은 단일특징 분할회귀트리를 사상한 네트워크에서의 학습결과, MRNN은 본 논문에서 제안한 네트워크에서의 학습결과를 의미한다. 도표의 각 수치는 위에서 정의된 MSE 를 측도로 하여 산출한 오차값이다.

도표 1-(a) 과 도표 2-(a)를 살펴보면, 분산 감소(variance reduction) 허용치를 5%, 10%, 20%로 나누어 실행한 결과 RT의 경우 트레이닝의 결과와 테스트의 결과가 10% 와 5% 에서 큰 간격이 벌어지고 모든 경우에서 RT보다 MRNN의 결과가 우수했음을 알 수 있다. 또한 실험 1과 실험 2, 각 경우에 MRNN의 결과가 SRNN보다 학습결과가 우수함을 알 수 있다.

실험 1과 실험 2를 통한 각 MRNN을 비교해 보면 RT의 결과에 의존적인 지도학습보다는 획득뉴런을 Gaussian basis 함수에 의해 조정된 경우의 학습결과가 비교적 향상되었음을 알 수 있다. 반면에, 두 실험에서 각 SRNN은 획득뉴런을 RT의 결과에 의존한 경우의 학습결과가 우수했다.

또한, 도표 1-(b), 도표 2-(b)를 통해서 과성장한 RT를 SRNN과 MRNN이 가지치기 했음을 알 수 있다. 즉 한번도 획득하지 못한 뉴런들과 그 링크들을 가지치기 함으로써 그 군집의 수가 감소하고 따라서 오차도 감소했음을 보여준다.

도표 1-(a) : 실험 1을 통한 MSE 회귀오차

% Variance Reduction To (User Specified)	RT	SRNN	MRNN
20 (training) (testing)	1.620 1.855	1.466 1.628	1.690 1.554
10	1.195 1.801	1.437 1.616	1.659 1.555
5	0.843 2.023	1.761 1.918	1.889 1.778

도표 1-(b) : 실험 1의 결과 산출된 군집의 수

% Variance Reduction To (User Specified)	RT	SRNN	MRNN
20	5	5	5
10	15	9	10
5	47	10	13

도표 2-(a) : 실험 2을 통한 MSE 회귀오차

% Variance Reduction To (User Specified)	RT	SRNN	MRNN
20 (training) (testing)	1.620 1.855	1.770 1.656	1.548 1.465
10	1.195 1.801	1.886 1.724	1.823 1.680
5	0.843 2.023	1.855 1.902	2.189 1.776

도표 2-(b) : 실험 2의 결과 산출된 군집의 수

% Variance Reduction To (User Specified)	RT	SRNN	MRNN
20	5	5	5
10	15	10	10
5	47	13	13

IV. 결 론

단일특징 분할 회귀트리를 계층신경망으로 전환한 기존의 기법이 초상자의 형태로 분할되는데 대해, 제안된 기법은 초다면체의 형태로 분할함으로써 보다 정확한 경계선을 도출, 학습성능이 향상되었다. 제안한 기법은 단일특징 분할 회귀트리를 다중특징 분할 회귀트리로 확장시킨 효과를 가지며 따라서 심화학습의 결과

가 우수하다. 또한 기존의 방법이 생성된 회귀트리의 결과에 의존적이기 때문에 두드러지지 않은 성능개선이 있었다면 본 연구는 회귀트리의 결과에 의존적인 경우에도 기존의 기법인 SRNN보다 학습능력이 우수했으며, 또한 획득뉴런을 Gaussian basis 함수를 이용하여 조정할 경우에는 더 큰 폭으로 SRNN보다 향상되었음을 관찰할 수 있었다.

참 고 문 헌

- [1] L. Brieman, J. Frieman, R. Olshen, and C.J. Stone, *Classification and Regression Trees*, Wadsworth Books, Belmont Calif., 1984.
- [2] I.K. Sethi, "Entropy nets: From decision tree to neural networks," *proc. IEEE*, Vol. 78, pp. 1605-1613, 1990.
- [3] I.K. Sethi and G.P.R. Sarvarayudu, "Hierarchical classifier design using mutual information," *IEEE Trans., PAMI*, Vol. 4, pp. 441-445, July, 1982.
- [4] J. Sietsma and R.J.F.Dow, "Neural Net Pruning-Why and How," *IEEE International Conference on Neural Networks*, San Diego, vol. 1, pp. 325-333, 1988.
- [5] M. Freat, "The Upstart Algorithm: A Method for Constructing and Training Feed forward Neural Networks," *Neural Computation* 2, pp. 198-209, 1990.
- [6] Moddy and C.J. Darken, "Fast learning in networks of locally tuned processing units," *Neural Computation*, Vol. 1, pp. 151-160, 1989.
- [7] D.F. Specht, "A general regression neural network," *IEEE Trans. Neural Networks*, Vol. 2, No. 6, pp. 568-576, 1991.
- [8] R.W. Prager, "CART / CMAC Hybrid: Regression trees with Interpolation," *proc. of 12th IAPR International Conference on Pattern Recognition*, Vol. II, Conf. B: Pattern Recognition and

- Neural Networks, pp. 476-478, 1994.
- [9] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973.
- [10] D.E. Rumelhart and J.L. McClelland, *Paralell Distributed Processing - Exploration in the Microstructure of Cognition Vol. 1: Foundations*, MIT Press, Cambridge, Mass., 1986
- [11] J. Hertz, A. Krogh and R. G. Palmer, *Introduction to the Theory of Neural Computation*, Addison Wesley, 1991.
- [12] I.K. Sethi and G. Yu, "A Neural Network Approach to Robot localization using Ultrasonic Sensors," proc. of 5th IEEE International Symposium on Intelligent Control, pp. 513-517, 1990.
- [13] J.S. Albus, "A new approach to manipulator control: cerebellar model articulation controller (CMAC)," ASME, Trans. G: Journal of Dynamic Systems, Measurements and Control, pp. 220-227, 1975.
- [14] S. Lim and S.C. Kim, "Design of fuzzy inference systems using regression trees," proc. of 1995 IEEE International Conference on Neural Networks, Vol. 1, pp. 345-348, 1995.

— 저 자 소 개 —



金 聖天(正會員)

1975년 서울대학교 공과대학
공업교육학(전기전공)(공학사).

1976년 ~ 1977년 동아컴퓨터
(주) Sys. Eng. 1977년 ~

1978년 스페리랜드 유니백,
Sales Rep. 1979년 Wayne

State Univ.(컴퓨터공학석사). 1982년 Wayne
State Univ. (컴퓨터 공학박사). 1982년 ~ 1984
년 캘리포니아 주립대학교 조교수. 1984년 ~ 1985
년 금성반도체(주) 책임연구원. 1990년 ~ 1991년
IEEE region 10 secretary. 1993년 ~ 현재 한
국정보과학회 병렬처리 시스템 연구회 위원장. 1985
년 ~ 현재 서강대학교 전자계산학과 교수로 재직.
관심분야는 병렬처리 시스템, Interconnection
Network, 신경 회로망, 이동통신, Computer
Network등임



林 淑(正會員)

1986년 이화여자대학교 자연대 수
학과(이학사). 1988년 서강대학교

전자계산학과(공학사). 1990년 서
강대학교 대학원 전자계산학과(공

학석사). 1991년 ~ 현재 서강대
학교 전자계산학과 박사과정. 1992

년 ~ 현재 국립여수수산대학교 조교수. 관심분야는 병
렬처리 시스템, 신경회로망등임