

연결선 파괴에 의한 인공 신경망의 크기 축소*

이재식** · 이혁주***

The Size Reduction of Artificial Neural Network by Destroying the Connections*

Jae Sik Lee** · Hyuk Ju Lee***

■ Abstract ■

A fully connected Artificial Neural Network (ANN) contains many connections. Compared to the pruned ANN with fewer connections, the fully connected ANN takes longer time to produce solutions and may not provide appropriate solutions to new unseen data. Therefore, by reducing the size of ANN, we can overcome the overfitting problem and increase the computing speed. In this research, we reduced the size of ANN by destroying the connections. In other words, we investigated the performance change of the reduced ANN by systematically destroying the connections. Then we found the acceptable level of connection-destruction on which the resulting ANN performs as well as the original fully connected ANN. In the previous researches on the size reduction of ANN, the reduced ANN had to be retrained every time some connections were eliminated. Therefore, it took long time to obtain the reduced ANN. In this research, however, we provide the acceptable level of connection-destruction according to the size of the fully connected ANN. Therefore, by applying the acceptable level of connection-destruction to the fully connected ANN without any retraining, the reduced ANN can be obtained efficiently.

Keyword : Artificial Neural Network, Size Reduction of Neural Network, Pruning of Neural Network, Connection-Destroyed Neural Network

논문접수일 : 2001년 3월 19일 논문게재확정일 : 2002년 2월 7일

* 이 논문은 2000년도 두뇌한국 21 사업에 의하여 지원되었음.

** 아주대학교 경영학부 교수

*** LG CNS e Solution사업부 CRM팀

1. 서론

산업이 점점 고도화되고, 인터넷 비즈니스와 같은 새로운 사업분야가 생김에 따라, 생성되는 데이터의 양은 엄청나게 늘어가고, 데이터로부터 얻을 수 있는 정보 또한 계속해서 늘어가고 있다. 우리는 이렇게 많은 데이터 속에 숨겨져 있는 수 많은 지식과 정보를 찾아내고자 통계적 방법이나 인공지능 기법 등을 이용하고 있는데, 매일 천문학적인 숫자로 늘어가고 있는 데이터를 분석하기란 여간 어려운 일이 아니다. 특히 인터넷 비즈니스를 위한 웹 로그 데이터 분석의 경우, 하루에 수 기가 바이트씩 쌓이는 데이터를 분석해야 하는 상황에 직면하게 될 수도 있다. 더구나 실시간으로 이런 엄청난 양의 데이터를 분석하고, 해답을 주고자 한다면, 데이터의 선별·정제·분류에 있어서 신중해야 할 뿐만 아니라, 분석을 위한 컴퓨팅 환경 역시 따라주어야 한다. 즉, 이런 분석을 할 때에 보다 적은 자원을 소모하여 만족할 만한 결과를 얻을 수 있다면, 대용량의 데이터 분석에 많은 도움이 될 수 있을 것이다. 본 연구에서는 이 점에 착안하여 인공 신경망(Artificial Neural Network)의 오류극복성(Fault Tolerance)을 이용해서 보다 빠른 속도로 만족할 만한 결과를 얻을 수 있는 인공 신경망의 개발 전략을 제시하고자 한다.

대부분의 인공 신경망 모델은 모든 노드를 완전 연결(Fully connected)하여 학습하는 방법을 쓰고 있는데, 이러한 방법은 노드 수가 수 만개 이상 되는 실제 문제를 풀 때에는 긴 계산 시간이 필요할 뿐만 아니라, 새로운 문제에 대해서는 적절한 해답을 제공하지 못하는 경우도 발생한다. 이런 단점을 해결하기 위하여는 인공 신경망의 복잡도(Complexity)를 줄여야 하는데, 연결선(Connection)들의 일부를 제거하거나, 은닉층의 처리요소(Processing Element) 즉 노드의 개수를 줄이는 방법을 쓸 수 있다. 본 연구에서는 완전 연결된 인공 신경망에서 작은 값의 가중치(Weight)를 가진 연결선을 제거해서 연결선의 개수를 줄이는 방법으로 인

공 신경망의 복잡도를 줄이고자 한다.

본 연구에서는 먼저 UCI(University of California at Irvine)의 기계학습 데이터 저장소[Blake *et al.*, 1998]로부터 데이터 집합을 몇 개 선정하여, 가능한 한 최상의 적중률이 나오도록 인공 신경망 모델들을 만든다. 다음, 각 모델마다 노드간 연결선을 가중치의 크기가 작은 것부터 계속해서 제거해 나가면서, 인공 신경망의 성능 저하 정도를 관찰한다. 연결선이 제거된 인공 신경망은 불완전하지만, 인공 신경망은 오류 극복성이 있을 뿐 만 아니라[Hansen *et al.*, 1991] 적중률에 큰 영향을 주지 않을 작은 값의 가중치를 가진 연결선부터 제거하기 때문에 어느 정도 수준까지는 인공 신경망의 성능이 크게 저하되지 않을 것이다. 마지막으로 실험 결과를 종합하여, 전체 연결선의 개수에 대해서 성능저하 없이 제거 가능한 연결선의 비율을 제시한다. 이 비율을 적용하여 연결선을 제거하면, 완전 연결된 인공 신경망과 성능이 비슷하며 컴퓨팅 속도가 빠른 인공 신경망을 얻을 수 있게 된다.

본 논문은 다음과 같이 구성되어 있다. 제 2절에서는 인공 신경망의 크기 축소에 관한 기존 연구에 대해 살펴보고, 제 3절에서는 인공 신경망의 연결선을 제거해 나가면서, 인공 신경망의 성능을 관찰·분석하고, 실험 결과를 바탕으로 연결선의 허용파괴수준을 제시한다. 제 4절에서는 검증용 데이터로 허용파괴수준에 대한 검증을 한다. 제 5절에서는 본 연구의 요약 및 의의, 그리고 한계점에 대해 기술한다.

2. 인공 신경망의 크기 축소에 관한 기존 연구

일반적으로 인공 신경망을 복잡하게 구축하면 주어진 문제의 Training 데이터는 잘 처리할 수 있지만 문제의 특성을 일반화시키는 데는 다소 어려움이 있다. 즉, 주어진 Training 데이터에 과잉 적합(Overfitting)할 가능성이 크므로 학습한 데이터 이

외의 데이터가 주어졌을 때 적절한 해답을 줄 수 없게 되는 것이다. 이는 Training 데이터에 노이즈가 포함되어 있는 경우에 본질적이지 못한 패턴까지도 학습하게 되기 때문이다. 이러한 문제점을 해결하기 위해서는 인공 신경망의 복잡도가 가능한 작은 것이 좋다. 인공 신경망의 크기를 축소시켜 복잡도를 줄이면, 학습 결과의 일반화가 용이하게 되며, 학습 속도를 향상시킬 수 있다[Chandrasekaran *et al.*, 2000].

인공 신경망의 크기 축소에 관한 연구로는 Mozer and Smolensky[1989], Sietsma and Dow[1991] 등의 인공 신경망의 노드의 제거에 관한 연구가 있고, Le Cun *et al.*[1989], Thodberg[1991], Ponnappalli *et al.*[1999] 등의 인공 신경망의 연결선의 제거에 관한 연구가 있다.

Mozer and Smolensky[1989]는 Skeletonisation이라는 기법을 이용해서 인공 신경망의 성능에 영향을 덜 미치는 노드를 제거함으로써, 인공 신경망의 크기를 축소시키는 방법을 연구하였다. Skeletonisation 기법은 다음과 같은 절차를 거쳐 수행된다. 먼저 1) 완전 연결된 인공 신경망을 적절한 성능에 도달할 때까지 훈련시키고, 2) 각각의 노드를 제거했을 때의 성능 변화를 계산한다. 다음, 3) 성능 변화에 가장 작은 영향을 미치는 노드를 제거하고, 4) 그 인공 신경망을 다시 훈련시킨다. 계속해서 미리 설정한 성능 수준에 도달할 때까지 2), 3), 4)의 과정을 반복한다. Mozer and Smolensky는 Skeletonisation 기법을 이용하여 크기를 축소 한 인공 신경망이 완전 연결된 인공 신경망보다 제한된 시간 안에 해의 발견능력이 더 우수하며, 동일한 해를 발견한다면 해를 발견하기까지의 속도가 더 빠르다고 주장하였다. 그들은 자동적으로 노드를 제거하는 방법을 제시하기는 했지만, 그 절차가 어느 시점에서 멈춰야 할 지에 대해서는 지침을 제시하지 못했다.

Sietsma and Dow[1991]는 성능에 큰 영향을 미치지 않는 비공헌 노드(Non-contributing nodes)와 불필요한 정보를 담고 있는 노드를 제거하는 방

법을 제시하였다. 비공헌 노드는 두 가지 종류의 노드로 정의되었는데, 그 중 한 가지는 일정 출력 노드(Constant output nodes)로서 입력값들에 대해 일정한 값을 출력하는 노드이며, 다른 한 가지는 중복 노드(Redundant nodes)로서 입력값들에 대해서 거의 같은 값을 출력하는 노드의 짝 중에 하나를 말한다.

Le Cun *et al.*[1989]은 인공 신경망의 성능에 영향을 덜 미치는 연결선을 제거하여 인공 신경망의 크기를 축소시키는 Optimal Brain Damage (OBD) 기법을 소개했다. OBD는 다음과 같은 절차를 거친다. 먼저 1) 완전 연결된 인공 신경망을 적절한 성능에 도달할 때까지 훈련시킨다. 다음, 2) 인공 신경망의 성능에 영향을 미치는 정보량을 의미하는 'Saliency'를 계산한다. OBD에서의 Saliency는 오차의 가중치에 대한 2차 도함수로 정의된다. 그리고 3) 계산된 Saliency 값들을 크기 순으로 정렬하여 낮은 Saliency를 갖는 연결선부터 n개의 연결선을 제거하고, 4) 그 인공 신경망을 다시 훈련시킨다. 계속해서 2), 3), 4)의 과정을 미리 설정한 적절한 성능에 도달할 때까지 반복한다.

Thodberg[1991] 역시 성능에 영향을 많이 미치지 않는 연결선을 제거하는 방식으로 인공 신경망의 크기를 축소시키는 연구를 했다. 그 과정을 보면 다음과 같다. 먼저 1) 완전 연결된 인공 신경망을 적절한 성능에 도달할 때까지 훈련시킨다. 2) 연결선 하나를 선택하여 제거하고, 3) 그 인공 신경망을 훈련시키고 성능을 평가한다. 다음, 4) 그 인공 신경망의 성능이 사전에 정해놓은 허용 가능한 오류 범위 내에 있으면, 그 인공 신경망을 유지시키고, 그렇지 않으면 원래 모습으로 다시 회복시킨다. 계속해서 2), 3), 4)의 과정을 각각의 연결선마다 수행한다.

Ponnappalli *et al.*[1999]은 LRSI(Local Relative Sensitivity Index)를 계산해서 LRSI가 작은 연결선을 제거하는 방법으로 인공 신경망의 크기를 축소시켰다. 여기서 'Sensitivity'란 개별 연결선이 제거되었을 때 인공 신경망의 성능에 영향을 미치는

정도를 말하며, 연결선이 제거될 때 SSE(Sum of squared errors)가 얼마만큼 변하는지를 계산해서 얻어진다. Ponnappalli *et al.*은 입력층과 은닉층 사이의 연결을 HIP(Hidden neuron to input layer)라 하고, 은닉층과 출력층 사이의 연결을 OHP(Output neuron to hidden layer)라 정의하였다. LRSI란 특정 HIP내의 모든 Sensitivity의 합에 대한 그 HIP에 속한 개별 연결선의 Sensitivity의 비율, 또는 OHP내의 모든 Sensitivity의 합에 대한 그 OHP에 속한 개별 연결선의 Sensitivity의 비율을 의미한다. Ponnappalli *et al.*이 제시한 연결선 제거 과정은 다음과 같다. 먼저 1) 완전 연결된 인공 신경망을 적절한 성능에 도달할 때까지 훈련시키고, 2) 각 연결선의 Sensitivity 값을 계산한다. 다음, 3) HIP와 OHP 각각에서 LRSI를 계산하고, 4) LRSI가 미리 정한 임계치 보다 작은 연결선들을 제거한다. 다음, 5) 연결선이 일부 제거된 인공 신경망을 다시 훈련 시키고, 6) 제거하기 전 인공 신경망과 그 성능을 비교한다. 미리 설정한 한계 SSE이내로 성능이 유지되는 동안 2)부터 6)까지의 과정을 반복하고, 한계를 넘으면, 인공 신경망의 축소를 중지한다.

3. 인공 신경망의 연결선 허용파괴수준

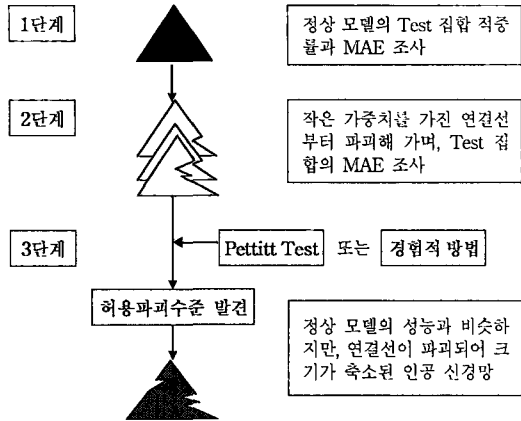
3.1 인공 신경망의 허용파괴수준 발견 단계

제 1절에서 언급했듯이 인공 신경망은 오류 극복성이 있기 때문에 완전 연결된 정상적인 인공 신경망의 구조가 일부 파괴되어도 성능이 급격하게 감소하지는 않는다. 특히 작은 가중치를 가진 연결선을 제거한다면, 큰 가중치를 가진 연결선을 제거할 때보다 성능이 덜 감소할 것이므로, 가중치가 작은 연결선들을 위주로 제거시킨다면 연결선 제거로 인한 성능 저하가 아주 천천히 일어날 것이다. 이러한 방식으로 연결선을 제거해 나가다가 성능이 급격히 감소하는 시점을 발견한다면, 그 이전

시점까지는 연결선을 제거해도 무방할 것이라는 추측을 할 수 있다.

본 연구에서는 전체 연결선 중 은닉층에서 출력층으로의 연결선을 제외한 나머지 연결선들을 ‘파괴대상 연결선’으로 선정하였다. 앞으로 ‘연결선의 제거’나 ‘연결선의 파괴’는 파괴대상 연결선의 제거나 파괴를 의미한다. 이렇게 선정한 이유는, 마지막 은닉층에서 출력층으로의 연결선이 우연히 많이 제거되면 인공 신경망의 출력에 문제가 발생하기 때문이다. 본 연구에서는 ‘연결선의 제거’보다는 ‘연결선의 파괴’라는 용어를 사용하고자 한다. 왜냐하면, 기존의 연구에서는 연결선을 제거하면서 계속적으로 축소된 인공 신경망을 다시 훈련시키는 반면, 본 연구에서는 완전 연결된 인공 신경망의 훈련이 일단 끝나면 다시는 훈련을 하지 않고 그냥 연결선을 없애버리기 때문이다. 파괴대상 연결선의 총 개수 중 파괴된 연결선의 개수를 파괴비율이라고 명명하였으며, 완전 연결된 인공 신경망과 거의 성능차이가 없지만 최대한 연결선을 파괴할 수 있는 파괴비율을 ‘허용파괴수준’으로 정의하였다. 이 허용파괴수준의 적용이 일반화 될 수 있다면, 기존의 연구들처럼 연결선을 제거할 때마다 다시 훈련을 반복해야 하는 번거로운 작업 없이, 완전 연결된 인공 신경망의 연결선들을 허용파괴수준까지 단 한번에 곧바로 파괴하여 사용할 수 있을 것이다.

본 연구의 과정은 크게 3개의 단계로 나뉘어 있다. 제 1단계에서는 완전 연결된 인공 신경망 모델을 만들어 Training 데이터 집합을 사용하여 훈련을 시키고, 제 2단계에서는 그 모델의 연결선들의 가중치를 모두 조사해 가중치가 작은 연결선부터 차례로 파괴하면서 Test 데이터 집합의 적중률과 MAE(Mean Absolute Error)의 추이를 본다. 제 3단계에서는 제 3.4절에서 소개할 Pettitt Test와 경험적 방법을 이용하여 Test 데이터 집합의 MAE 증가 변화점을 찾는다. 최종적으로 이 변화점을 분석하여 적절한 허용파괴수준을 제시한다. 이러한 연구 단계를 도식화하면 [그림 3-1]과 같다.



[그림 3-1] 허용파괴수준 발견 단계

본 연구에서 사용된 데이터들은 목표 변수가 이산형 변수들이다. 그러므로, 제 2단계에서 연결선 파괴 후 적중률을 계산할 때에 인공 신경망의 출력치가 적중했는지 아닌지를 판정하기 위한 임계치를 임의로 정해야 한다. 그런데 임계치의 설정에 따라 적중률의 추이가 약간씩 변화되기 때문에 적중률은 허용파괴수준을 정하는 판단근거로는 부적절했다. 따라서 본 연구에서는 [그림 3-1]의 제 2단계에서 보는 바와 같이 적중률을 관찰하지는 않고, 단지 MAE만을 관찰하여 허용파괴수준을 제시하였다.

3.2 정상 인공 신경망 모델 구축

3.2.1 모델집합(Model Set) 선정

허용파괴수준을 발견하기 위해 사용할 데이터로는

UCI의 자료 중 Credit Screening과 Breast Cancer를 선정하였다. Credit Screening은 Credit Card 신청의 수락여부를 결정하는 문제로서 총 690개의 인스턴스를 가지고 있고, Breast Cancer는 유방암의 재발여부를 판정하는 문제로서 총 286개의 인스턴스를 가지고 있다. 이 데이터들을 Training 집합, Test 집합, Evaluation 집합으로 나눌 때에 집합간의 비율을 6 : 3 : 1과 5 : 4 : 1의 두 종류로 하였다. 또한, 각 비율로 모델집합을 만들 때에 각 모델집합의 데이터 구성이 다양하게 되도록 하기 위해서 원본 데이터로부터 세 번 샘플링(Sampling)을 하였다. 즉, Credit Screening의 경우 6 : 3 : 1의 비율로 샘플링을 세 번하고, 또 5 : 4 : 1의 비율로 샘플링을 세 번하여 총 6 종류의 모델집합을 만드는 것이다. 같은 방법으로 Breast Cancer의 경우에도 6 종류의 모델집합을 만들어, 총 12개의 모델집합을 만들었다. 이 12개의 각 모델집합마다 은닉층의 개수를 1개, 2개, 3개로 변화시켜서 <표 3-1>과 같이 총 36개의 모델을 만들었다. 이와 같이 동일한 데이터에 대해서 은닉층의 개수를 다르게 한 이유는 여러 복잡도를 가진 인공 신경망을 구축하기 위해서이다.

Credit Screening과 Breast Cancer는 모두 이진분류 문제이다. 입력값은 0에서 1사이로 정규화하였고, 목표값은 0 또는 1로 설정하였다. 인공 신경망의 출력값이 0.35이하일 때는 0으로 0.65 이상일 때는 1로 정하였으며 0.35와 0.65 사이의 값이

<표 3-1> 실험을 위한 모델의 설정

| 데이터 집합 비율 | | Train : Test : Eval ¹⁾ - 6 : 3 : 1 | | | Train : Test : Eval ¹⁾ - 5 : 4 : 1 | | |
|------------------|--------|---|--------|--------|---|--------|--------|
| | | 1 | 2 | 3 | 1 | 2 | 3 |
| Sampling 번호 | | | | | | | |
| Credit Screening | 은닉층 1개 | 모델 I_1 | 모델 I_2 | 모델 I_3 | 모델 I_4 | 모델 I_5 | 모델 I_6 |
| | 은닉층 2개 | II_1 | II_2 | II_3 | II_4 | II_5 | II_6 |
| | 은닉층 3개 | III_1 | III_2 | III_3 | III_4 | III_5 | III_6 |
| Breast Cancer | 은닉층 1개 | IV_1 | IV_2 | IV_3 | IV_4 | IV_5 | IV_6 |
| | 은닉층 2개 | V_1 | V_2 | V_3 | V_4 | V_5 | V_6 |
| | 은닉층 3개 | VI_1 | VI_2 | VI_3 | VI_4 | VI_5 | VI_6 |

주) 1) Training 집합 : Test 집합 : Evaluation 집합의 비율

출력되었을 때에는 맞추지 못한 것으로 하였다. 이 임계치는 제 1단계의 훈련시에만 사용하는 것이고, 제 2단계에서는 전술한 바와 같이 MAE를 사용하여 성능을 측정하였다.

3.2.2 정상 모델 구축

각 모델은 적중률이 가장 높도록 은닉노드의 개수를 조정하여 <표 3-2>에서 보는 바와 같이 Credit Screening에 대한 모델 18개와 Breast Cancer

<표 3-2> 정상 모델의 연결선의 개수와 적중률

| 모델 | 모멘텀 | 학습율 | I 노드 ¹⁾ | H 노드 ²⁾ | 연결선 ³⁾ | 적중률 ⁴⁾ | MAE ⁴⁾ | Eval ⁵⁾ |
|-------|-----|-----|--------------------|--------------------|-------------------|-------------------|-------------------|--------------------|
| I_1 | 0.4 | 0.5 | 30 | H1-25 | 750 | 84.57 | 0.191 | 62.86 |
| I_2 | 0.4 | 0.5 | 30 | H1-5 | 150 | 85.57 | 0.184 | 84.29 |
| I_3 | 0.4 | 0.5 | 30 | H1-15 | 450 | 85.07 | 0.185 | 85.71 |
| I_4 | 0.4 | 0.5 | 30 | H1-20 | 600 | 88.17 | 0.177 | 87.14 |
| I_5 | 0.4 | 0.5 | 30 | H1-25 | 750 | 86.26 | 0.186 | 85.71 |
| I_6 | 0.4 | 0.5 | 30 | H1-5 | 150 | 85.88 | 0.169 | 84.29 |
| II_1 | 0.4 | 0.5 | 30 | H1-15, H2-5 | 525 | 85.57 | 0.170 | 60 |
| II_2 | 0.4 | 0.5 | 30 | H1-20, H2-1 | 620 | 87.06 | 0.165 | 85.71 |
| II_3 | 0.4 | 0.5 | 30 | H1-20, H2-3 | 660 | 85.88 | 0.176 | 85.71 |
| II_4 | 0.4 | 0.5 | 30 | H1-15, H2-9 | 585 | 87.40 | 0.174 | 88.57 |
| II_5 | 0.4 | 0.5 | 30 | H1-20, H2-9 | 780 | 85.88 | 0.169 | 87.14 |
| II_6 | 0.4 | 0.5 | 30 | H1-15, H2-9 | 585 | 87.02 | 0.169 | 84.29 |
| III_1 | 0.4 | 0.5 | 30 | H1-20, H2-9, H3-4 | 816 | 86.06 | 0.165 | 61.43 |
| III_2 | 0.4 | 0.5 | 30 | H1-15, H2-15, H3-5 | 750 | 86.57 | 0.173 | 85.71 |
| III_3 | 0.4 | 0.5 | 30 | H1-20, H2-5, H3-4 | 720 | 84.08 | 0.165 | 88.57 |
| III_4 | 0.4 | 0.5 | 30 | H1-20, H2-15, H3-4 | 960 | 86.26 | 0.165 | 87.14 |
| III_5 | 0.4 | 0.5 | 30 | H1-20, H2-15, H3-6 | 990 | 86.26 | 0.159 | 85.71 |
| III_6 | 0.4 | 0.5 | 30 | H1-20, H2-15, H3-6 | 990 | 86.64 | 0.154 | 87.14 |
| IV_1 | 0.6 | 0.5 | 10 | H1-12 | 120 | 97.62 | 0.052 | 97.14 |
| IV_2 | 0.4 | 0.5 | 10 | H1-6 | 60 | 93.81 | 0.083 | 95.71 |
| IV_3 | 0.4 | 0.5 | 10 | H1-4 | 40 | 93.81 | 0.068 | 92.86 |
| IV_4 | 0.4 | 0.5 | 10 | H1-2 | 20 | 95.70 | 0.063 | 91.43 |
| IV_5 | 0.4 | 0.5 | 10 | H1-8 | 80 | 95.34 | 0.070 | 95.71 |
| IV_6 | 0.4 | 0.5 | 10 | H1-2 | 20 | 94.00 | 0.083 | 95.71 |
| V_1 | 0.4 | 0.5 | 10 | H1-12, H2-6 | 192 | 97.14 | 0.051 | 98.57 |
| V_2 | 0.4 | 0.5 | 10 | H1-6, H2-4 | 84 | 94.29 | 0.074 | 95.71 |
| V_3 | 0.4 | 0.5 | 10 | H1-4, H2-4 | 56 | 95.24 | 0.062 | 94.29 |
| V_4 | 0.4 | 0.5 | 10 | H1-4, H2-5 | 112 | 96.06 | 0.060 | 91.43 |
| V_5 | 0.4 | 0.5 | 10 | H1-4, H2-8 | 72 | 95.70 | 0.067 | 95.71 |
| V_6 | 0.4 | 0.5 | 10 | H1-10, H2-6 | 160 | 93.91 | 0.082 | 97.14 |
| VI_1 | 0.9 | 0.6 | 10 | H1-12, H2-6, H3-3 | 210 | 98.57 | 0.032 | 95.71 |
| VI_2 | 0.4 | 0.5 | 10 | H1-10, H2-6, H3-3 | 178 | 95.24 | 0.066 | 95.71 |
| VI_3 | 0.9 | 0.6 | 10 | H1-10, H2-4, H3-3 | 152 | 95.24 | 0.057 | 95.71 |
| VI_4 | 0.4 | 0.5 | 10 | H1-10, H2-6, H3-3 | 178 | 96.24 | 0.049 | 94.29 |
| VI_5 | 0.4 | 0.5 | 10 | H1-6, H2-4, H3-3 | 96 | 95.70 | 0.054 | 94.29 |
| VI_6 | 0.4 | 0.5 | 10 | H1-6, H2-4, H3-3 | 96 | 94.27 | 0.077 | 94.29 |

주) 1) 입력층 노드의 개수

2) 은닉층 노드의 개수로서 H1은 첫 번째 은닉층, H2는 두 번째 은닉층, H3는 세 번째 은닉층 노드의 개수

3) 마지막 은닉층에서 출력층으로의 연결선을 제외한 총 연결선의 개수, 즉 파괴대상 연결선의 개수

4) Test 집합의 적중률과 MAE이고 적중률의 단위는 %

5) Evaluation 집합의 적중률이고 단위는 %

에 대한 모델 18개를 구축하였다. 예를 들어, <표 3-2>에서 Credit Screening에 대한 모델들인 모델 I_1, 모델 I_2, 모델 I_3을 보면, 은닉층의 개수는 1이고 은닉노드의 개수는 각각 25, 5, 15이다. 여기서 은닉층의 개수 1은 임의로 정한 것이지만, 은닉노드의 개수는 적중률이 가장 높게 나오도록 실험에 의해서 정해진 것이다. 이 세 모델은 <표 3-1>에서 보는 바와 같이 Training 집합, Test 집합, Evaluation 집합의 비율이 6:3:1로 동일하지만, 각 데이터 집합을 구성하고 있는 데이터들이 다르다. 즉, 이 세 모델은 비록 은닉층의 개수가 1개로 동일하지만, 다른 내용의 데이터로 훈련을 함으로써 은닉노드의 개수가 달라진 것이다.

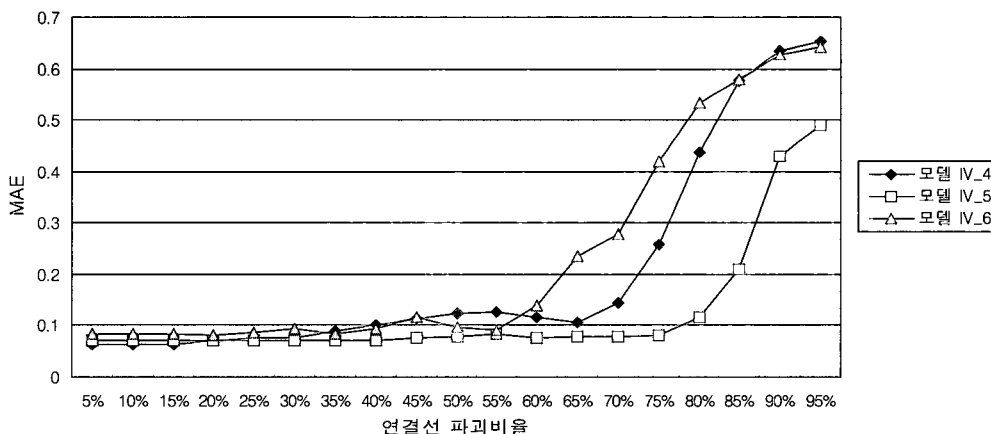
36개 모델의 훈련에는 모두 백프로퍼게이션 학습 알고리즘을 사용했으며, 전이함수는 시그모이드 함수(Sigmoid Function)를 사용하였다. 실험을 통해 Credit Screening의 경우 학습 회수 5만번, Breast Cancer의 경우 학습 회수 10만번일 때 RMSE(Root Mean Squared Error)가 더 이상 떨어지지 않는 것을 관찰할 수 있었기 때문에 학습회수를 Credit Screening은 5만번, Breast Cancer는 10만번으로 정하였다. 모멘텀(Momentum)과 학습율(Learning Rate)은 모델 대부분 0.4와 0.5로 고정시켰다. 결과적으로, 각 모델의 연결선의 개수와 적중률이 <표 3-2>와 같이 도출되었다.

3.3 인공 신경망의 연결선 파괴 실험

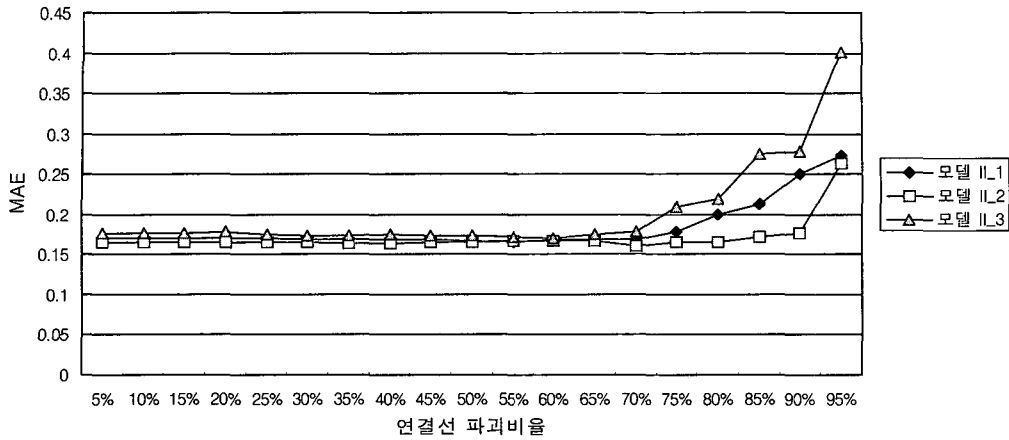
각 모델별로 연결선의 가중치를 조사하여, 작은 가중치를 가진 연결선부터 파괴하면서 인공 신경망의 MAE가 어떻게 변화하는가를 살펴보았다. [그림 3-2]는 36개의 모델 가운데 모델 IV_4, 모델 IV_5, 모델 IV_6의 연결선의 파괴비율 증가에 따른 MAE의 증가 추세를 보여주고 있는데, 인공 신경망의 성능에 거의 변화가 없다가 어떤 파괴비율 이상이 되면, 성능이 크게 저하되고 있음을 알 수 있다.

연결선의 일부를 파괴하는 실험을 통해 인공 신경망은 어느 수준까지는 거의 성능이 변하지 않는다는 것을 알 수 있다. 부록의 [그림 A-1]~[그림 A-10]에서 볼 수 있듯이 모든 모델의 MAE는 적어도 파괴비율 15%까지는 정상 모델의 MAE와 비교해 크게 차이 나지 않는다는 것을 알 수 있다. 하지만, 모델별로 성능 저하 시점이 다르기 때문에 허용파괴수준을 일반화하기에 어려움이 있다.

본 연구에서 다루는 인공 신경망 모델들의 연결선 개수는 <표 3-2>에서 보듯이, 가장 적은 것은 20개이고, 가장 많은 것은 990개이다. 이 모델들을 연결선 개수 450개를 기준으로 두 군집으로 나누었다. 즉 연결선 개수가 450개 미만인 것은 상대적으로 복잡도가 낮은 것으로서 <범주 1>로 정의했으며, 450개 이상의 것은 상대적으로 복잡도가 높



[그림 3-2] 파괴비율 별 MAE 추이(모델 IV_4, 모델 IV_5, 모델 IV_6)



[그림 3-3] 파괴비율 별 MAE 추이(모델 II_1, 모델 II_2, 모델 II_3)

은 것으로 간주하여 <범주 2>로 정의했다. <범주 2>의 모델들 중에는 연결선이 60% 이상 파괴되어도 성능 저하가 일어나지 않는 모델도 있다. [그림 3-3]은 <범주 2>에 속하는 모델인 모델 II_1, 모델 II_2, 모델 II_3의 MAE 변화 추이인데, [그림 3-2]의 <범주 1> 모델에 비해 성능이 쉽게 저하되지 않음을 알 수 있다. 따라서 복잡도가 높은 모델에서는 복잡도가 낮은 모델보다 허용파괴수준을 높게 설정할 수 있을 것이다.

3.4 허용파괴수준 조사

[그림 3-2]와 [그림 3-3]에서 보듯이, 파괴비율을 높이면 인공 신경망 모델의 성능이 대부분 일정 수준이 지나면서 크게 떨어지는 것을 알 수 있다. 즉, 성능 저하 패턴이 어떤 점을 기준으로 변한다고 가정 할 수 있다. 예를 들어 [그림 3-2]의 모델 IV_4를 보면, 파괴 비율이 약 35%에서 45%로 넘어가면서 성능이 급격히 떨어지기 시작하는 것을 볼 수 있다. 그러므로, 35%~45% 부근에 성능 변화 패턴을 구분할 수 있는 변화점이 있다고 할 수 있을 것이다. 즉, 그 변화점을 기준으로 해서 그 이전의 성능 저하 모습은 거의 변화없는 패턴을 가지고 있고, 그 이후로는 성능 저하 속도가 급격히 올라가는 패턴을 가지고 있다고 할 수 있다. 우

리는 이러한 점을 데이터 분포의 구조적 변화점을 발견하는 Pettitt Test[Pettitt, 1979]로 찾을 수 있다고 판단하였다.

Pettitt Test는 비모수적 변화점 발견 기법(Non-parametric Change-point Detection Method)으로서 시간 순차적 데이터(Time Sequence Data)의 분포에서 구조적으로 변화가 있는 부분을 찾아내는 방법이다[Oh and Han, 2000]. 분포에 구조적 변화가 있다면, Pettitt Test를 이용해 찾은 변화점을 전 후로 해서 변화점 이전의 분포와 이후의 분포는 다른 패턴을 갖고 있는 셈이 된다. Pettitt Test는 다음과 같이 변화점을 찾아낸다. 먼저, 각 관측치 X_1, X_2, \dots, X_N 에 순위를 1부터 N까지 매긴다. 이때 r_i 값은 X_i 의 순위가 된다. 다음 j 지점에서 W_j 를 다음과 같이 계산한다.

$$W_j = \sum_{i=1}^j r_i, \quad j = 1, 2, \dots, N-1 \quad (1)$$

다음 매 지점 j 마다 $K_{m,n}$ 을 구한다.

$$K_{m,n} = \max_{j=1, 2, \dots, N-1} |2W_j - j(N+1)|, \quad (2)$$

등식 (2)에서 얻은 $K_{m,n}$ 값에 해당하는 지점 j 가 변화점이 되며, m 으로 표시한다. n 은 관측치의 수 N 에서 m 을 뺀 값으로 변화점 이후의 관측치의

수가 된다. 변화점은 X_1, X_2, \dots, X_N 을 m 과 n 개의 두 부분으로 분리하는데, 변화점 이전의 m 개에 해당하는 관측치와 변화점 이후의 n 개에 해당하는 관측치는 서로 다른 패턴을 가지고 있다고 할 수 있다.

우리는 Pettitt Test를 통해 제시된 파괴비율이 허용파괴수준이 될 가능성이 있다고 보고, 그 파괴 비율을 'P허용파괴수준'이라고 명명하였다. <표 3-2>의 36개 모델을 <범주 1>과 <범주 2>로 구분한 후, Pettitt Test를 한 결과는 <표 3-3>과 같다.

<표 3-3> P허용파괴수준을 이용한 파괴 실험 결과

| 범 주 | 모 델 | 연결선 ¹⁾ | 파괴전 MAE ²⁾ | P허용파괴수준 | 파괴후 MAE ³⁾ | MAE 증가 ⁴⁾ |
|-------|-------|-------------------|-----------------------|---------|-----------------------|----------------------|
| 1 | IV_4 | 20 | 0.063 | 40.00% | 0.101 | + 0.038 |
| | IV_6 | 20 | 0.083 | 55.00% | 0.092 | + 0.008 |
| | IV_3 | 40 | 0.068 | 50.00% | 0.074 | + 0.006 |
| | V_3 | 56 | 0.062 | 46.43% | 0.074 | + 0.012 |
| | IV_2 | 60 | 0.083 | 50.00% | 0.088 | + 0.005 |
| | V_5 | 72 | 0.067 | 44.44% | 0.139 | + 0.072 |
| | IV_5 | 80 | 0.070 | 45.00% | 0.076 | + 0.006 |
| | V_2 | 84 | 0.074 | 45.24% | 0.111 | + 0.037 |
| | VI_5 | 96 | 0.054 | 45.83% | 0.108 | + 0.054 |
| | VI_6 | 96 | 0.077 | 45.83% | 0.122 | + 0.045 |
| | V_4 | 112 | 0.060 | 44.64% | 0.083 | + 0.023 |
| | IV_1 | 120 | 0.052 | 45.83% | 0.058 | + 0.006 |
| | I_2 | 150 | 0.184 | 34.67% | 0.196 | + 0.012 |
| | I_6 | 150 | 0.169 | 48.67% | 0.177 | + 0.008 |
| | VI_3 | 152 | 0.057 | 46.05% | 0.125 | + 0.068 |
| | V_6 | 160 | 0.082 | 40.63% | 0.095 | + 0.013 |
| | VI_2 | 178 | 0.066 | 44.94% | 0.102 | + 0.036 |
| | VI_4 | 178 | 0.049 | 44.94% | 0.111 | + 0.062 |
| V_1 | 192 | 0.051 | 39.06% | 0.053 | + 0.002 | |
| VI_1 | 210 | 0.032 | 40.48% | 0.064 | + 0.032 | |
| 2 | I_3 | 450 | 0.185 | 65.33% | 0.184 | - 0.001 |
| | II_1 | 525 | 0.170 | 68.57% | 0.169 | - 0.001 |
| | II_4 | 585 | 0.174 | 41.03% | 0.181 | + 0.007 |
| | II_6 | 585 | 0.169 | 47.86% | 0.170 | + 0.001 |
| | I_4 | 600 | 0.177 | 43.33% | 0.179 | + 0.002 |
| | II_2 | 620 | 0.165 | 48.39% | 0.165 | 0.000 |
| | II_3 | 660 | 0.176 | 60.61% | 0.170 | - 0.006 |
| | III_3 | 720 | 0.165 | 44.44% | 0.171 | + 0.006 |
| | I_1 | 750 | 0.191 | 56.00% | 0.188 | - 0.003 |
| | I_5 | 750 | 0.186 | 56.00% | 0.186 | 0.000 |
| | III_2 | 750 | 0.173 | 61.33% | 0.173 | 0.000 |
| | II_5 | 780 | 0.169 | 29.36% | 0.168 | - 0.001 |
| | III_1 | 816 | 0.165 | 46.57% | 0.174 | + 0.009 |
| | III_4 | 960 | 0.165 | 45.83% | 0.167 | + 0.002 |
| III_5 | 990 | 0.159 | 44.44% | 0.163 | + 0.004 | |
| III_6 | 990 | 0.154 | 44.44% | 0.159 | + 0.005 | |

주) 1) 파괴대상 연결선의 개수

2) 완전 연결된 인공 신경망 모델의 Test 집합의 MAE

3) P허용파괴수준까지 파괴했을 때의 인공 신경망의 Test 집합의 MAE

4) (파괴후 MAE) - (파괴전 MAE)

<표 3-3>에서 완전 연결된 인공 신경망의 MAE와 P허용과피수준까지 연결선을 파괴한 인공 신경망의 MAE를 보면, <범주 2> 모델의 경우에는 MAE간에 차이가 거의 없다는 것을 알 수 있다. 하지만, <범주 1> 모델들은 <범주 2> 모델들에 비해 MAE가 크게 증가하였다. 가장 심한 것은 모델 V_5인데, MAE가 0.067에서 0.139로 약 2배 정도 증가하였다. 이와 같은 현상은, <범주 1> 모델의 경우에 연결선의 개수가 너무 적어서 중요한 가중치를 가진 연결선이 <범주 2> 모델보다 쉽게 파괴되었기 때문이라고 판단된다. 따라서, <범주 1> 모델에 대해서 P허용과피수준을 그대로 적용하는 것은 무리가 있다. 즉, <범주 1> 모델의 허용과피수준은 P허용과피수준보다는 보수적으로 정해져야 할 필요가 있다.

<범주 1> 모델에 대해서는 다음과 같은 휴리스틱 방법으로 새로운 허용과피수준을 구하였다.

- 1) 연결선 1개를 파괴하였을 때의 MAE를 MAE_1 이라 하고, P허용과피수준에서의 MAE를 MAE_p 라고 하고, P허용과피수준 직전의 MAE를 MAE_{p-1} 이라 하자.
- 2) MAE_1 으로부터 MAE_{p-1} 까지의 평균을 구하고, 이를 MAE_A 라 하자.
- 3) MAE_1 과 MAE_p 사이에 있고 MAE_A 보다 작은 MAE들 중에서 MAE_p 에 가장 가까운 MAE를 찾고, 이를 MAE_M 이라 하자.
- 4) MAE_M 일 때의 파괴 비율을 D_M 이라 하자.

이렇게 구한 파괴비율 D_M 이 우리가 새롭게 사용할 평균(Mean) 허용과피수준으로서 'M허용과피수준'이라고 명명하였다.

<표 3-4>는 연결선이 72개인 <범주 1>에 해당하는 모델 V_5의 연결선 파괴 실험의 결과표로서 이 표를 보고 위의 과정을 적용해 보자. 이 모델의 P허용과피수준은 44.44%이고 MAE_p 는 0.139이다. 먼저, MAE_1 은 0.067이고, MAE_{p-1} 는 0.126이다. MAE_1 0.067로부터 MAE_{p-1} 0.126까지의 평

균 MAE_A 를 구하면 0.075가 된다. 다음, MAE_1 0.067과 MAE_p 0.139 사이에 있고, MAE_A 0.075보다 작은 MAE들 중에서 MAE_p 0.139에 가장 가까운 MAE인 MAE_M 은 0.074가 된다. 이때의 파괴 비율인 D_M 은 26.39%이고, 이것이 바로 모델 V_5의 M허용과피수준이 된다.

<표 3-4> 모델 V_5의 연결선 파괴 실험 결과

| 파괴 연결선 수 | MAE | 파괴비율 | 파괴 연결선 수 | MAE | 파괴비율 |
|----------|--------------|--------------|----------|--------------|---------------|
| 1 | 0.067 | 1.39% | 14 | 0.068 | 19.44% |
| 2 | 0.067 | 2.78% | 15 | 0.067 | 20.83% |
| 3 | 0.067 | 4.17% | 16 | 0.067 | 22.22% |
| 4 | 0.067 | 5.56% | 17 | 0.070 | 23.61% |
| 5 | 0.067 | 6.94% | 18 | 0.072 | 25.00% |
| 6 | 0.067 | 8.33% | 19 | 0.074 | 26.39% |
| 7 | 0.069 | 9.72% | 20 | 0.078 | 27.78% |
| 8 | 0.068 | 11.11% | 22 | 0.089 | 30.56% |
| 9 | 0.068 | 12.50% | 24 | 0.090 | 33.33% |
| 10 | 0.068 | 13.89% | 26 | 0.103 | 36.11% |
| 11 | 0.068 | 15.28% | 28 | 0.105 | 38.89% |
| 12 | 0.068 | 16.67% | 30 | 0.126 | 41.67% |
| 13 | 0.068 | 18.06% | 32 | 0.139 | 44.44% |

<표 3-5>는 <범주 1> 모델들에 M허용과피수준을 적용했을 때의 MAE를 보여주고 있다. 이 표에서 보듯이, <범주 1>의 모든 모델들에서 완전 연결된 인공 신경망의 MAE와 M허용과피수준까지 연결선을 파괴한 인공 신경망의 MAE의 차이가 미미하다는 것을 알 수 있다.

지금까지의 실험 결과를 보면 각 범주마다 다른 허용과피수준을 적용하는 것이 좋다는 결론을 내릴 수 있다. <표 3-3>과 <표 3-5>에서 <범주 1>을 보면 P허용과피수준을 적용할 때 보다 M허용과피수준을 적용할 때 MAE의 증가가 더 작다는 것을 알 수 있다. 그러므로, 안전한 연결선 파괴를 위해서 M허용과피수준을 <범주 1>의 허용과피수준으로 정하였다. <범주 2>의 경우에는 <표 3-3>에서 보듯이 P허용과피수준을 적용해도 MAE

<표 3-5> M허용파괴수준을 이용한 파괴 실험 결과

| 범 주 | 모 델 | 연결선 ¹⁾ | 파괴전 MAE ²⁾ | M허용파괴수준 | 파괴후 MAE ³⁾ | MAE 증가 ⁴⁾ |
|------|------|-------------------|-----------------------|---------|-----------------------|----------------------|
| 1 | IV_4 | 20 | 0.063 | 20.00% | 0.070 | + 0.007 |
| | IV_6 | 20 | 0.083 | 35.00% | 0.083 | 0.000 |
| | IV_3 | 40 | 0.068 | 47.50% | 0.073 | + 0.005 |
| | V_3 | 56 | 0.062 | 28.57% | 0.064 | + 0.002 |
| | IV_2 | 60 | 0.083 | 33.33% | 0.084 | + 0.001 |
| | V_5 | 72 | 0.067 | 26.39% | 0.074 | + 0.007 |
| | IV_5 | 80 | 0.070 | 40.00% | 0.071 | + 0.001 |
| | V_2 | 84 | 0.074 | 26.19% | 0.076 | + 0.002 |
| | VI_5 | 96 | 0.054 | 31.25% | 0.062 | + 0.008 |
| | VI_6 | 96 | 0.077 | 31.25% | 0.086 | + 0.009 |
| | V_4 | 112 | 0.060 | 21.43% | 0.061 | + 0.001 |
| | IV_1 | 120 | 0.052 | 35.00% | 0.052 | 0.000 |
| | I_2 | 150 | 0.184 | 21.33% | 0.187 | + 0.003 |
| | I_6 | 150 | 0.169 | 26.00% | 0.169 | 0.000 |
| | VI_3 | 152 | 0.057 | 19.74% | 0.062 | + 0.005 |
| | V_6 | 160 | 0.082 | 28.13% | 0.083 | + 0.001 |
| | VI_2 | 178 | 0.066 | 28.09% | 0.069 | + 0.003 |
| | VI_4 | 178 | 0.049 | 22.47% | 0.052 | + 0.003 |
| V_1 | 192 | 0.051 | 33.85% | 0.050 | - 0.001 | |
| VI_1 | 210 | 0.032 | 19.05% | 0.034 | + 0.002 | |

주) 1) 파괴대상 연결선의 개수

2) 완전 연결된 인공 신경망 모델의 Test 집합의 MAE

3) M허용파괴수준까지 파괴했을 때의 인공 신경망의 Test 집합의 MAE

4) (파괴후 MAE) - (파괴전 MAE)

의 증가가 크지 않기 때문에 P허용파괴수준을 <범 주 2>의 허용파괴수준으로 정하였다. 하지만, 각 범 주에 속하는 여러 인공 신경망 모델들 각각에 대해서 상이한 값의 허용파괴수준을 제시할 수는 없다. 즉, 각 범주를 대표하는 하나의 허용파괴수준을 정해야 하는데, 본 연구에서는 보수적인 입장을 취하여 각 범주의 허용파괴수준들 중에서 가장 작은 값을 채택하였다. <표 3-5>를 보면 M허용파괴수준의 값들 중 가장 작은 값이 모델 VI_1일 때의 19.05%이므로 <범 주 1>의 허용파괴수준은 19%

로 결정하였고, <표 3-3>의 <범 주 2>를 보면 P허용파괴수준의 값들 중 가장 작은 값이 모델 II_5일 때의 29.36%이므로 <범 주 2>의 허용파괴수준은 29%로 결정하였다.

<표 3-6>은 <범 주 1>과 <범 주 2> 각각에 대해서 결정된 허용파괴수준을 적용했을 때의 각 모델 별 MAE를 보여주고 있는데, 표에서 보듯이 모든 모델의 MAE는 완전 연결된 인공 신경망의 MAE와 비교해 큰 차이가 없음을 알 수 있다.

〈표 3-6〉 허용파괴수준 적용 결과

| 범 주 | 모 델 | 연결선 ¹⁾ | 파괴전 MAE ²⁾ | 허용파괴수준 ³⁾ | 파괴후 MAE ⁴⁾ | MAE 증가 ⁵⁾ |
|-------|-------|-------------------|-----------------------|----------------------|-----------------------|----------------------|
| 1 | IV_4 | 20 | 0.063 | 20% | 0.070 | + 0.007 |
| | IV_6 | 20 | 0.083 | 20% | 0.082 | - 0.001 |
| | IV_3 | 40 | 0.063 | 20% | 0.070 | + 0.007 |
| | V_3 | 56 | 0.062 | 19.64% | 0.064 | + 0.002 |
| | IV_2 | 60 | 0.083 | 20% | 0.083 | 0.000 |
| | V_5 | 72 | 0.067 | 19.44% | 0.068 | + 0.001 |
| | IV_5 | 80 | 0.070 | 20% | 0.070 | 0.000 |
| | V_2 | 84 | 0.074 | 19.05% | 0.075 | + 0.001 |
| | VI_5 | 96 | 0.054 | 19.79% | 0.061 | + 0.007 |
| | VI_6 | 96 | 0.077 | 19.79% | 0.081 | + 0.004 |
| | V_4 | 112 | 0.060 | 19.64% | 0.061 | + 0.001 |
| | IV_1 | 120 | 0.052 | 19.17% | 0.052 | 0.000 |
| | I_2 | 150 | 0.184 | 19.33% | 0.184 | 0.000 |
| | I_6 | 150 | 0.169 | 19.33% | 0.169 | 0.000 |
| | VI_3 | 152 | 0.057 | 19.08% | 0.062 | + 0.005 |
| | V_6 | 160 | 0.082 | 19.38% | 0.082 | 0.000 |
| | VI_2 | 178 | 0.066 | 19.10% | 0.066 | 0.000 |
| VI_4 | 178 | 0.049 | 19.10% | 0.051 | + 0.002 | |
| V_1 | 192 | 0.051 | 19.27% | 0.054 | + 0.003 | |
| VI_1 | 210 | 0.032 | 19.05% | 0.034 | + 0.002 | |
| 2 | I_3 | 450 | 0.185 | 29.78% | 0.183 | - 0.002 |
| | II_1 | 525 | 0.170 | 29.14% | 0.169 | - 0.001 |
| | II_4 | 585 | 0.174 | 30.77% | 0.165 | - 0.009 |
| | II_6 | 585 | 0.169 | 32.48% | 0.169 | 0.000 |
| | I_4 | 600 | 0.177 | 30.00% | 0.176 | - 0.001 |
| | II_2 | 620 | 0.165 | 29.03% | 0.165 | 0.000 |
| | II_3 | 660 | 0.176 | 31.82% | 0.173 | - 0.003 |
| | III_3 | 720 | 0.165 | 31.94% | 0.167 | + 0.002 |
| | I_1 | 750 | 0.191 | 30.67% | 0.190 | - 0.001 |
| | I_5 | 750 | 0.186 | 29.33% | 0.187 | + 0.001 |
| | III_2 | 750 | 0.173 | 30.67% | 0.171 | - 0.002 |
| | II_5 | 780 | 0.169 | 29.36% | 0.168 | - 0.001 |
| | III_1 | 816 | 0.165 | 31.86% | 0.167 | + 0.002 |
| | III_4 | 960 | 0.165 | 30.21% | 0.164 | - 0.001 |
| | III_5 | 990 | 0.159 | 29.29% | 0.160 | + 0.001 |
| III_6 | 990 | 0.154 | 29.29% | 0.153 | - 0.001 | |

주) 1) 파괴대상 연결선의 개수

2) 완전 연결된 인공 신경망 모델의 Test 집합의 MAE

3) 파괴되는 연결선의 개수는 정수이므로 정확히 19%와 29%를 허용파괴수준으로 할 수는 없고 19%와 29%의 근사치를 허용파괴수준으로 적용하였음.

4) 허용파괴수준까지 파괴했을 때의 인공 신경망의 Test 집합의 MAE

5) (파괴후 MAE) - (파괴전 MAE)

4. 허용파괴수준 검증

검증을 위한 데이터로는 UCI 데이터 중에서 6개를 선정하여 사용하였다. <범주 1>에 대한 검증용 데이터로는 Pima Indians Diabetes(Indians Diabetes로 칭하기로 함), BUPA Liver Disorders(Liver Disorders로 칭하기로 함), Heart Disease를 사용하였다. Indians Diabetes는 피마족 인디언의 21세 이상 여성들 중에서 당뇨병 보유여부를 판정하는 문제로서 총 768개의 인스턴스를 가지고 있고, Liver Disorders는 환자의 간장에 이상이 있는가를 판정하는 문제로서 총 345개의 인스턴스를 가지고 있고, Heart Disease는 환자가 심장병을 가지고 있나를 판정하는 문제로서 총 270개의 인스턴스를 가지고 있다. <범주 2>에 대한 검증용 데이터로는 Ionosphere, Sonar, Dermatology를 사용하였다. Ionosphere는 성층권의 상부인 전리층에 자유전자가 존재하는지를 판정하는 문제로서 총 351개의 인스턴스를 가지고 있고, Sonar는 수중 음파 탐지기의 신호를 반사한 물체가 금속인가 암석인가를 판정하는 문제로서 총 208개의 인스턴스를 가지고 있고, Dermatology는 피부병의 종류를 진단하는 문제로서 총 366개의 인스턴스를 가지고 있다.

이들 데이터로부터 Training 집합, Test 집합, Evaluation 집합을 부작위로 6 : 3 : 1의 비율로 추

출하였으며, 검증을 위한 인공 신경망 모델은 모두 학습 알고리즘은 백프로퍼게이션을, 전이함수는 시그모이드 함수를 사용하였다. Indians Diabetes의 경우 입력변수가 총 8개이므로 입력층의 노드 개수 역시 8개이다. 은닉층은 모두 3개로 설정하였는데, 각 층의 노드개수는 실험에 의해서 6개, 4개, 2개로 구해졌다. 출력층의 노드 개수는 1개이므로 총 연결선의 개수가 82개가 되지만, 마지막 은닉층과 출력층간의 연결선 2개를 제외하면 파괴 대상 연결선의 개수는 총 80개가 된다. 모멘텀은 0.6, 학습율은 0.6으로 정해졌으며, 학습 회수는 10만번까지 하였다. 모멘텀과 학습율은 Test 집합의 MAE가 가장 낮을 때의 조합으로 구해졌다. 이와 같은 방법으로 검증용 모델들을 설정한 결과는 <표 4-1>과 같다.

이렇게 모델을 구축한 후 각각의 모델을 범주에 따라 해당 허용파괴수준을 적용하여 인공 신경망을 파괴하였다. 즉, Indians Diabetes, Liver Disorders, Heart Disease는 파괴 대상 연결선의 개수가 450개 미만으로서 <범주 1>에 해당하므로 M허용파괴수준인 19%를 적용하여 인공 신경망을 파괴하였고, Ionosphere, Sonar, Dermatology는 파괴 대상 연결선의 개수가 450개 이상으로서 <범주 2>에 해당하므로 P허용파괴수준인 29%를 적용하여 인공 신경망을 파괴하였다. 그 결과는 <표 4-2>와 같다.

<표 4-1> 검증을 위한 모델의 설정

| 범주 | 모델 | 모멘텀 | 학습율 | I 노드 ¹⁾ | H 노드 ²⁾ | 연결선 ³⁾ | 학습회수 |
|----|------------------|-----|-----|--------------------|--------------------|-------------------|------|
| 1 | Indians Diabetes | 0.6 | 0.6 | 8 | H1-6, H2-4, H3-2 | 80 | 10만번 |
| | Liver Disorders | 0.7 | 0.7 | 6 | H1-4 | 24 | 10만번 |
| | Heart Disease | 0.3 | 0.7 | 13 | H1-8 | 104 | 10만번 |
| 2 | Ionosphere | 0.3 | 0.7 | 34 | H1-23 | 782 | 10만번 |
| | Sonar | 0.7 | 0.7 | 60 | H1-15 | 900 | 10만번 |
| | Dermatology | 0.8 | 0.6 | 34 | H1-22, H2-11 | 990 | 10만번 |

주) 1) 입력층 노드의 개수

2) 은닉층 노드의 개수로 H1은 첫 번째 은닉층, H2는 두 번째 은닉층, H3는 세 번째 은닉층 노드의 개수

3) 마지막 은닉층에서 출력층으로의 연결선을 제외한 총 연결선의 개수, 즉 파괴대상 연결선의 개수

〈표 4-2〉 허용파괴수준 검증 결과

| 범주 | 모델 | 연결선 ¹⁾ | 파괴전 MAE ²⁾ | 허용파괴수준 | 파괴된 연결선 ³⁾ | 파괴후 MAE ⁴⁾ | MAE의 증가 ⁵⁾ |
|----|------------------|-------------------|-----------------------|--------|-----------------------|-----------------------|-----------------------|
| 1 | Indians Diabetes | 80개 | 0.330 | 18.75% | 15개 | 0.336 | + 0.006 |
| | Liver Disorders | 24개 | 0.354 | 20.83% | 5개 | 0.358 | + 0.004 |
| | Heart Disease | 104개 | 0.234 | 19.23% | 20개 | 0.233 | - 0.001 |
| 2 | Ionosphere | 782개 | 0.123 | 29.03% | 227개 | 0.116 | - 0.007 |
| | Sonar | 900개 | 0.241 | 29.00% | 261개 | 0.235 | - 0.006 |
| | Dermatology | 990개 | 0.205 | 29.09% | 288개 | 0.195 | - 0.010 |

주) 1) 마지막 은닉층에서 출력층으로의 연결선을 제외한 총 연결선의 개수, 즉 파괴대상 연결선의 개수

2) 완전 연결된 인공 신경망의 Evaluation 집합의 MAE

3) 파괴대상 연결선 중에서 허용파괴수준을 적용할 때 파괴된 연결선의 개수

4) 허용파괴수준으로 연결선을 파괴한 인공 신경망의 Evaluation 집합의 MAE

5) (파괴후 MAE) - (파괴전 MAE)

〈표 4-2〉에서 보듯이, 〈범주 1〉에 속하는 인공 신경망이건 〈범주 2〉에 속하는 인공 신경망이건 간에 해당 허용파괴수준까지 연결선을 파괴한 결과를 보면, MAE에 미세한 변화를 보일 뿐이다. 예를 들어, 〈범주 1〉에 속하는 모델인 Indians Diabetes의 경우에, 연결선을 파괴하기 전에 MAE는 0.330이었는데, 〈범주 1〉의 허용파괴수준 19%의 근사치인 18.75%까지 연결선을 파괴하였을 때 MAE는 0.336이 되었다. 즉, MAE가 0.006만큼 아주 미미하게 증가한 것이다. 같은 범주에 속하는 Liver Disorders도 역시 MAE가 0.004만큼 미미하게 증가하였으며, Heart Disease는 오히려 MAE가 0.001만큼 감소하였다. 한편 〈범주 2〉에 속하는 모델에서는 해당 허용파괴수준인 29%까지 연결선을 파괴한 후에 MAE가 모두 감소하였다.

검증 결과를 보면, 본 연구에서 제시하는 허용파괴수준까지 연결선을 파괴한 인공 신경망의 성능이 연결선 파괴전의 성능과 별 차이가 없음을 알 수 있다. 따라서, 완전 연결된 인공 신경망을 구축한 후에 연결선의 개수가 450개 미만인 경우에는 가중치가 작은 연결선부터 19%까지 파괴하고, 연결선의 개수가 450개 이상 990개 이하인 경우에는 가중치가 작은 연결선부터 29%까지 파괴해도 성

능에 큰 변화 없이 빠른 속도의 인공 신경망을 사용할 수 있다는 결론을 내릴 수 있다.

5. 결 론

본 연구에서는 인공 신경망의 특징인 오류 극복성을 근거로 하여 의도적으로 인공 신경망을 파괴하여 크기를 축소함으로써, 성능은 완전 연결된 인공 신경망과 비슷하지만 속도는 더 빠른 인공 신경망을 만드는 방법을 제시하였다. 즉, 인공 신경망의 연결선들을 어느 수준 만큼 파괴해도 그 성능이 떨어지지 않는지를 관찰하여, 성능이 떨어지기 전 적절한 시점을 찾아내서 그 시점까지 인공 신경망의 연결선을 파괴하여 크기를 축소하였다.

인공 신경망의 크기를 축소하는 방법에 관한 기존 연구의 경우, 연결선을 제거하면서 그 때의 성능을 측정하고, 목표한 성능에 도달할 때까지 계속해서 훈련 과정을 반복한다. 이러한 방법은 연결선 제거 후의 훈련 과정이 목표한 성능에 도달할 때까지 계속 반복되기 때문에 연결선의 가중치가 매 훈련 때마다 계속해서 바뀌게 될 뿐만 아니라 반복되는 훈련 과정만큼 인공 신경망의 크기 축소에 많은 시간이 걸리게 된다. 그러나 본 연구에서는, 실험

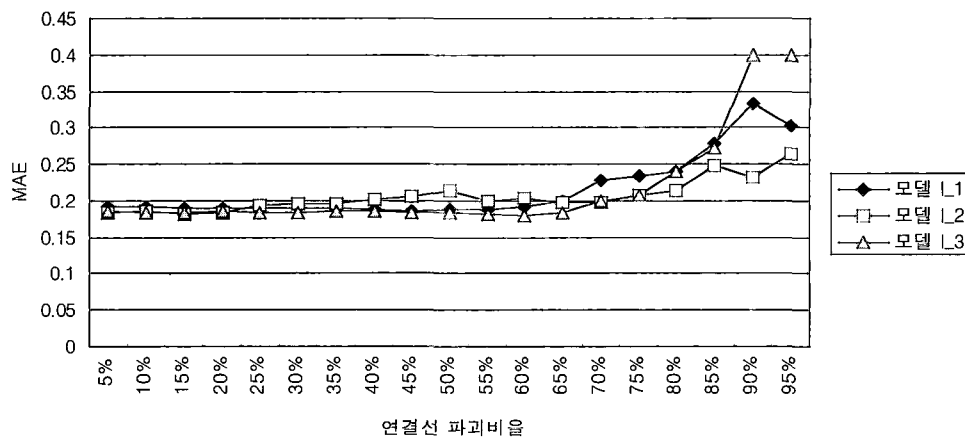
결과 얻은 허용파괴수준을 완전 연결된 인공 신경망에 아무런 훈련과정 없이 그대로 적용함으로써 훨씬 효율적으로 크기가 축소된 인공 신경망을 얻을 수 있도록 하였다. 즉, 완전 연결된 새로운 인공 신경망의 크기를 축소할 때에, 기존의 연구에서는 연결선 제거후의 학습 과정을 매번 다시 반복해야 하지만, 본 연구의 방법을 따를 경우에는 구축 완료된 인공 신경망의 연결선들을 이미 경험적으로 얻어낸 허용파괴수준까지 단 한번에 파괴함으로써 인공 신경망의 크기 축소가 끝나기 때문에 훨씬 시간이 적게 걸리고 효율적이다.

본 연구의 가장 큰 한계점은 연구 결과 얻어진 허용파괴수준의 일반화를 검증하기가 쉽지 않다는 것이다. 허용파괴수준을 하나의 공식으로 도출할 수는 없었으므로, 본 연구에서는 다양한 데이터에 본 연구결과를 적용·검증함으로써 일반화에 접근하는 방법을 채택하였다. 허용파괴수준을 정할 때에, 연결선의 개수에 따라 인공 신경망 모델의 범주를 나누고, 그 범주를 대표하는 허용파괴수준을 정하였는데, 이런 방법은 연결선의 개수가 그 범주를 벗어나는 경우에는 허용파괴수준을 제시할 수가 없게 된다. 또한, 각 범주에 대하여 허용파괴수

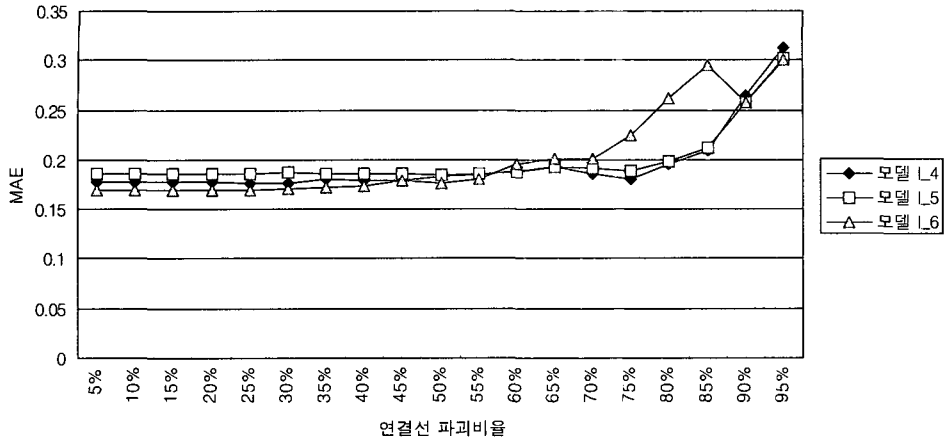
준을 정하는 방법도 단순히 범주 안의 모델들이 가지는 허용파괴수준들 중에서 가장 작은 것을 허용파괴수준으로 정하는 보수적인 방법을 채택하였는데, 이 방법도 과학적 근거를 제시할 수 없는 실험적 결과에 불과하다. 하지만 이러한 한계점들은 좀더 여러 가지 데이터를 사용하여 다양한 개수의 연결선으로 구성된 인공 신경망들을 가지고 실험을 한다면 해소될 수 있으리라고 판단된다.

추가로 언급할 수 있는 한계점은 인공 신경망 구축시에 입력변수 선정에 관한 문제를 고려하지 않았다는 것이다. 입력변수를 적절히 선정함으로써 인공 신경망의 크기를 축소할 수가 있다. 입력변수 선정의 문제는 그 자체가 큰 연구주제이며 많은 학자들이 연구하고 있다[John *et al.*, 1994 ; Setiono and Liu, 1997 ; 이재식과 차봉근, 1999]. 본 연구에서는 입력변수 선정문제는 연구 범위 밖으로 설정하고, 인공 신경망의 크기 축소 문제만 다루었다. 하지만, 향후에는 입력변수가 최적으로 선정되어 구축된 인공신경망의 연결선을 어느 정도 파괴하여도 성능의 변화가 없는지에 대한 연구가 수행되어야 할 것이다.

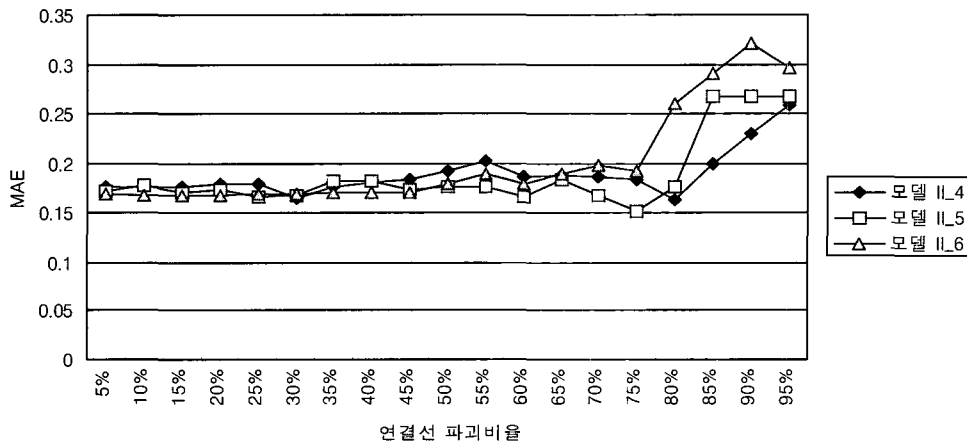
<부록>



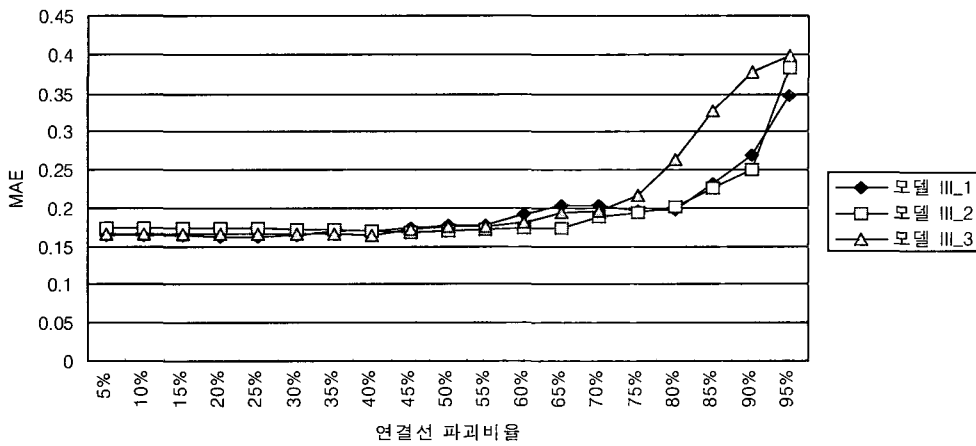
[그림 A-1] 파괴비율 별 MAE 추이(모델 L1, 모델 L2, 모델 L3)



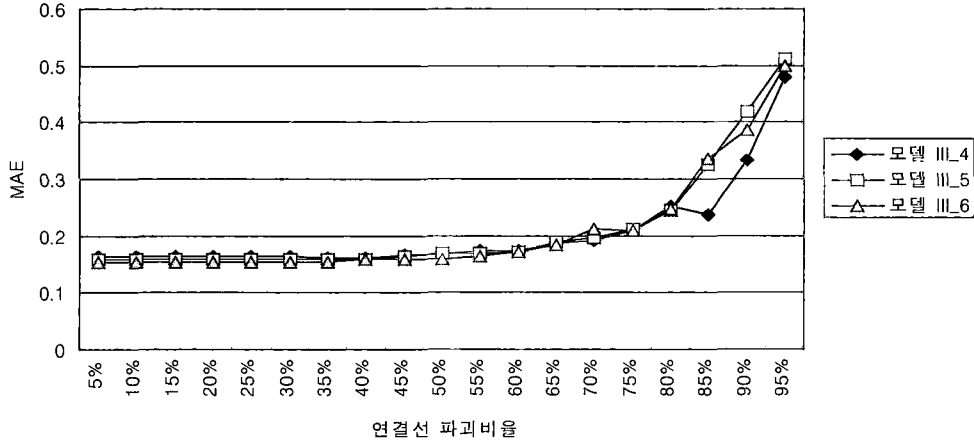
[그림 A-2] 파괴비율 별 MAE 추이(모델 I_4, 모델 I_5, 모델 I_6)



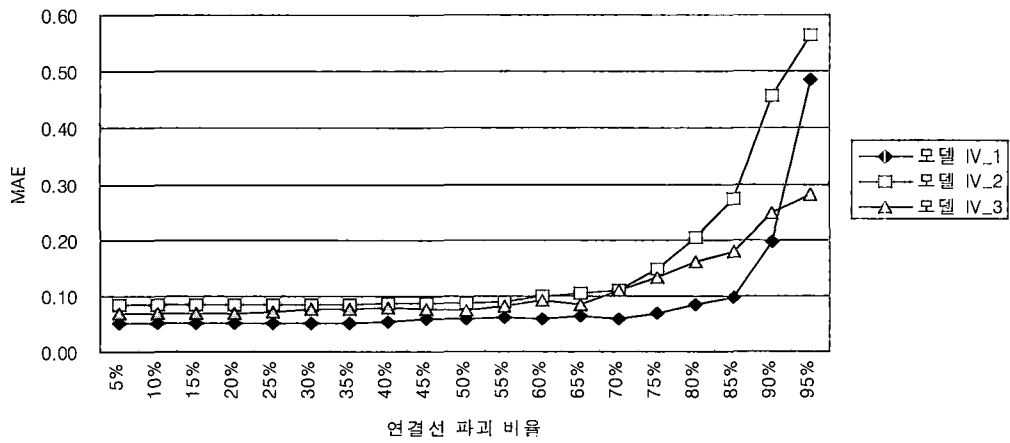
[그림 A-3] 파괴비율 별 MAE 추이(모델 II_4, 모델 II_5, 모델 II_6)



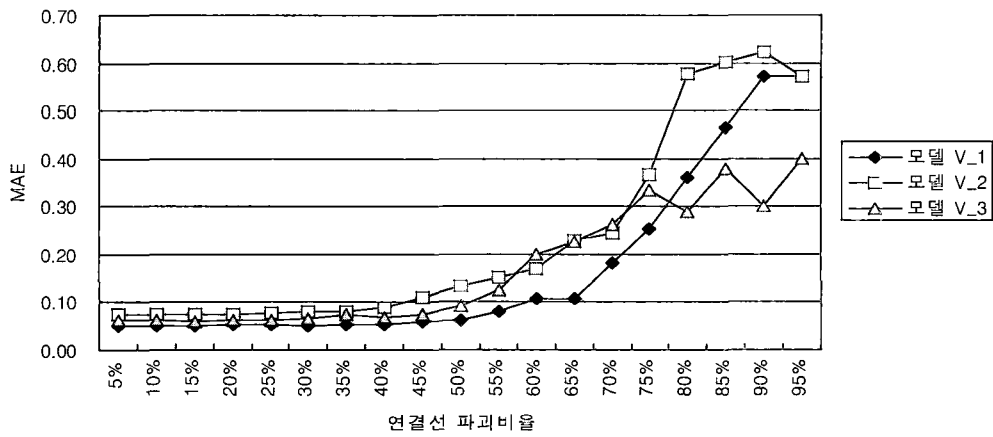
[그림 A-4] 파괴비율 별 MAE 추이(모델 III_1, 모델 III_2, 모델 III_3)



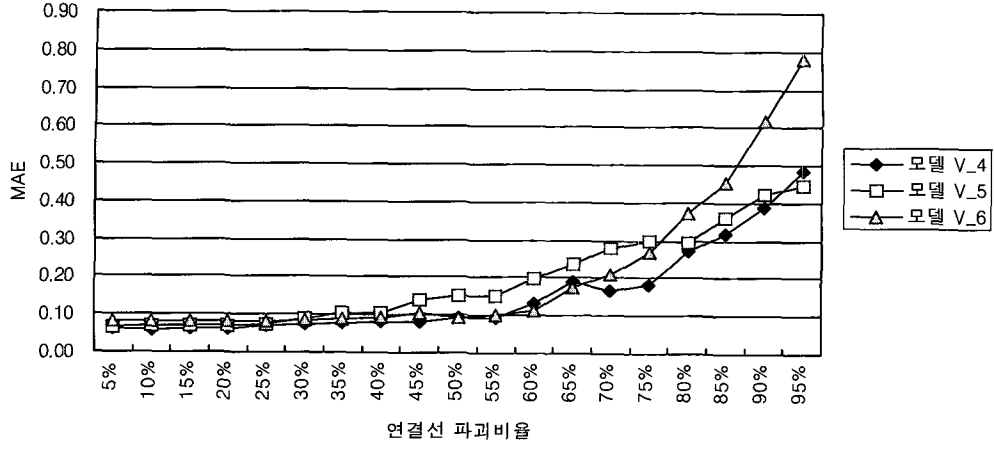
[그림 A-5] 파괴비율 별 MAE 추이(모델 III_4, 모델 III_5, 모델 III_6)



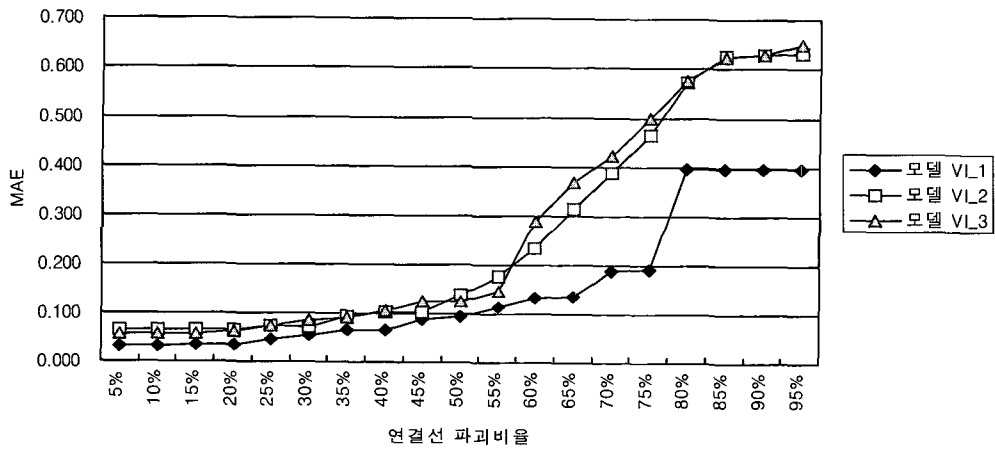
[그림 A-6] 파괴비율 별 MAE 추이(모델 IV_1, 모델 IV_2, 모델 IV_3)



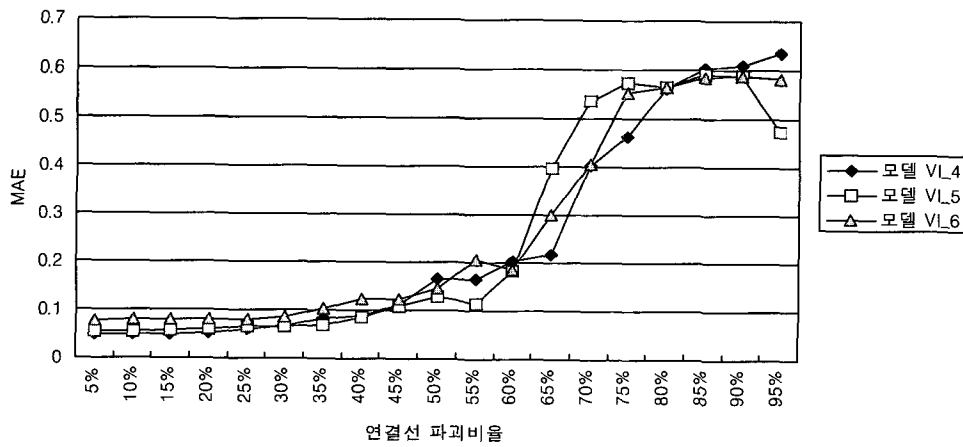
[그림 A-7] 파괴비율 별 MAE 추이(모델 V_1, 모델 V_2, 모델 V_3)



[그림 A-8] 파괴비율 별 MAE 추이(모델 V_4, 모델 V_5, 모델 V_6)



[그림 A-9] 파괴비율 별 MAE 추이(모델 VI_1, 모델 VI_2, 모델 VI_3)



[그림 A-10] 파괴비율 별 MAE 추이(모델 VI_4, 모델 VI_5, 모델 VI_6)

참고 문헌

- [1] 이재식, 차봉근, “유전적 알고리즘을 이용한 인공신경망의 구조 설계”, 「한국경영과학회지」, 제24권, 제3호(1999), pp.49-62.
- [2] Blake, C., E. Keogh and C.J. Merz, UCI Repository of Machine Learning Databases, [http://www.ics.uci.edu/~mlearn/MLRepository.html], Irvine, CA : Univ. of CA, Dept. of Information and Computer Science, 1998.
- [3] Chandrasekaran, H., H.H. Chen and M.T. Manry, “Pruning of Basis Functions in Nonlinear Approximators,” *Neurocomputing* (2000), pp.29-53.
- [4] Hansen, J.V., F. William and J.R. Messier, “Artificial Neural Networks : Foundations and Application to Decision Problem,” *Expert Systems with Applications*, 3(1991), pp.135-141.
- [5] John, G.H., R. Kohavi and K. Pfleger, “Irrelevant Features and the Subset Selection Problem,” in Cohen, W.W. and H. Hirsh (eds.), *Machine Learning : Proc. of the 11th Int'l Conf.*, Morgan Kaufmann Pub., San Francisco, CA(1994), pp.121-129.
- [6] Le Cun, Y., J.S. Denker and S.A. Solla, “Optimal Brain Damage,” in *Advances in Neural Information Processing (2)*, D.S. Touretzky (ed), (1989), pp.598-605.
- [7] Mozer, M.C. and P. Smolensky, “Using Relevance to Reduce Network Size Automatically,” *Connection Science*, 1(1989), pp. 3-16.
- [8] Oh, K.J. and I. Han, “Using Change-point Detection to Support Artificial Neural Networks for Interest Rates Forecasting,” *Expert Systems with Applications*, 19(2000), pp.105-115.
- [9] Pettitt, A.N., “A Non-parametric Approach to the Change-point Problem,” *Applied Statistics*, 28(2), (1979), pp.126-135.
- [10] Ponnappalli, P.V.S., K.C. Ho and M. Thomson, “A Formal Selection and Pruning Algorithm for Feedforward Artificial Neural Network Optimization,” *IEEE Trans. on Neural Networks*, 10(4), (1999), pp.964-968.
- [11] Setiono, R. and H. Liu, “Neural-network Feature Selector,” *IEEE Trans. on Neural Networks*, 8(3), (1997), pp.654-662.
- [12] Sietsma, J. and R.J.F. Dow, “Creating Artificial Neural Networks that Generalize,” *Neural Networks*, 4(1991), pp.67-79.
- [13] Thodberg, H.H., “Improving Generalization of Neural Networks through Pruning,” *Int'l Journal of Neural Systems*, 1(1991), pp. 317-326.