

# 유전적 알고리즘을 이용한 인공신경망의 구조 설계

이재식\* · 차봉근\*\*

## Structure Design of Artificial Neural Networks using Genetic Algorithm

Jae Sik Lee\* · Bong-Geun Cha\*\*

### ■ Abstract ■

Artificial Neural Networks(ANN) have been successfully applied to various kinds of business and engineering problems, especially those involved in pattern classification. However, because of the lack of design standard or guidelines, the structure of specific ANN depends on the designer's own experiments or choices. In other words, even though we could construct a better ANN, we often settled down with just a satisfactory ANN. The purpose of this research is to apply the Genetic Algorithm(GA) to design a structure of ANN that yields better performance compared to the existing test results. For a bankruptcy prediction problem, an existing research using ANN which consists of 22 input processing elements(PEs) for financial ratios and 5 hidden PEs showed 70% hit ratio. In our research, the input financial ratios and the number of hidden PEs are determined by GA. The best ANN, which consists of 8 input PEs and 6 hidden PEs, shows 78.03% hit ratio. In addition, we compare the performance of two types of reproduction schemes, i.e., generational reproduction and steady-state reproduction.

## 1. 서론

인공신경망은 전통적 통계 기법보다 우수한 예측력을 보이며, 규칙기반 전문가시스템의 한계

점인 지식획득 문제를 완화시켜줄 수 있다는 장점 때문에 문제해결의 규칙을 명시적으로 추출할 수 없는 문제에 많이 적용되어 왔다[Nelson and Illingworth, 1991]. 경영 문제의 경우에는 투자

\* 아주대학교 경영대학 경영학부 교수

\*\* IBM Korea 보험산업서비스 사업부

의사결정[Kamijo and Tanigawa, 1990], 기업 도산 예측[Odom and Sharda, 1990; 이재식과 한재홍, 1995], 재무 위험 관리[Bansal *et al.*, 1993], 기업신용평가[한인구 등, 1994] 등에 적용되었다. 인공신경망의 설계과정에서는 입력변수, 은닉층의 개수, 은닉층 처리요소 개수, 전이함수, 학습방법 등 여러 가지 결정하여야 할 요인들이 있다. 인공신경망의 설계에 관한 많은 연구와 실험들이 있었지만 아직까지 정립된 이론이나 기준이 없기 때문에 설계시 여러 요인들의 결정은 설계자의 경험에 의존해야 한다. 특히, 입력변수의 경우에는 그 수가 많아지게 되면 학습 시간이 오래 걸리게 되며, 만일 입력변수들 사이에 잡음이 섞여 있다면 학습 직중률이 저하되는 결과를 가져오기도 한다. 그러므로 적절한 입력변수들을 선정하는 것은 직중률과 학습 효율을 높이는 데 중요한 요소가 된다.

한편, 유전적 알고리즘[Goldberg, 1989]도 전통적 방법의 한계를 극복하는 장점을 갖고 있기 때문에 여러 분야에서 해를 찾는 도구로 이용되어 왔다. 유전적 알고리즘은 문제영역에서 병렬적으로 해를 탐색하기 때문에 매우 효율적인데, 이러한 장점은 인공신경망의 구조를 효율적으로 설계해보고자 하는 연구 분야에도 활용되어 많은 연구결과가 발표되고 있다[Opitz and Shavlik, 1994; Medsker, 1995; Back *et al.*, 1996; Yang and Honavar, 1997].

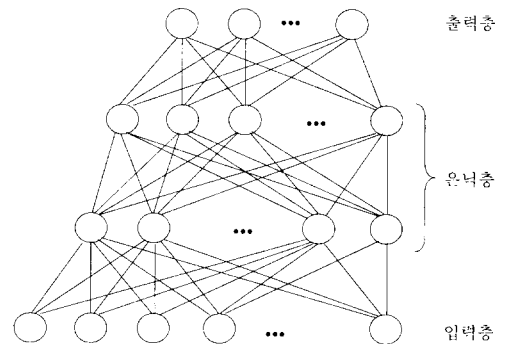
본 연구에서는 기업의 도산예측을 위한 인공신경망의 설계에 초점을 맞추고자 하는데, 인공신경망의 구조 설계시 고려할 수 있는 여러 요인들 중에서 어떠한 입력변수를 선정하고 은닉층의 처리요소 개수를 몇 개로 하였을 때 인공신경망의 직중률이 향상되는지를 연구하고자 한다. 이러한 문제의 해를 찾는 도구로서 유전적 알고리즘을 이용하였다. 본 논문은 다음과 같이 구성되어 있다. 제2절에서는 인공신경망의 기본 구조 및 설계시 고려하여야 할 요인들에 대하여 살펴본다. 제3절에서는 본 연구의 적용기법인 유전적 알고리즘에 대하여 용어 및 특징을 간략히 기술하고, 제4절에서는

인공신경망 구조설계에 관한 기존 연구들을 소개한다. 제5절에서는 본 연구에서 제시하는 모델인 ANN\_D\_GA의 구조에 대하여 설명하고, 제6절은 본 연구의 핵심부분으로서, 본 연구에서 수행한 실험의 설계 및 그 결과를 기술한다. 제7절에서는 본 연구에서 얻은 결론, 한계점 및 향후 연구과제 등을 제시한다.

## 2. 인공신경망

### 2.1 인공신경망 설계시의 고려 사항

인공신경망은 생물학에서 연구된 동물 뇌의 내부 구조를 모방하여 정보를 처리하는 방법으로서, 그 일반적인 구조는 [그림 2-1]과 같다.



[그림 2-1] 인공신경망의 구조

[그림 2-1]에서 보듯이 인공신경망은 입력층, 은닉층, 그리고 출력층으로 구성되어있고 각 층은 처리요소(processing elements : PE) 또는 노드(node)들로 구성되어 있다. 각 처리요소는 입력값에 대하여 전이함수를 통하여 계산된 출력값을 내보낸다. 그리고 각 층 사이의 처리요소들을 연결하는 가지(arc)는 연결강도(weight)를 가지고 있다. 인공신경망을 설계할 때에는 다음과 같은 것들을 고려해야 한다.

#### ① 학습 자료와 테스트 자료의 크기

학습과 테스트를 하기 위해서 어느 정도의 자

료가 필요한지, 또한 학습 자료와 테스트 자료는 어떠한 비율로 할 것인가를 결정하여야 한다. 학습 자료가 너무 많게 되면 과잉적합(overfitting)의 문제가 발생하게 되고 너무 적으면 만족할만한 테스트 결과를 주지 못한다.

② 학습 방법

각 가지의 연결강도를 조정하는 방법으로서 흔히 사용되는 학습방법은 백프로퍼게이션(back-propagation)방법이다. 인공신경망의 용도 또는 적용문제의 특성에 따라 다양한 방법이 연구되어 있다.

③ 토폴로지(Topology)

토폴로지란 각 층의 처리요소 개수 및 은닉층의 수를 말하는데, 이는 인공신경망의 학습효율뿐만 아니라 테스트 결과에도 중요한 영향을 미친다. 설계자가 마련한 입력변수의 후보들 중에서 어떤 것들을 선정할 것인가, 은닉층의 수 및 그 안의 처리요소 개수는 몇 개로 할 것인가에 대한 결정에는 아직 표준적인 방법이 없다.

④ 전이함수(Transfer Function)

은닉층과 출력층의 처리요소들은 입력된 값에 대하여 전이함수를 적용하여 출력값을 산출한다. 주로 비선형 함수들이 사용되는데, 이에선 계단함수(hard limiter), 임계논리(threshold logic) 함수, 그리고 시그모이드(sigmoid) 함수 등이 있다. 어떠한 전이함수를 사용할 것인가는 문제 영역에 따라 달라질 수 있으며 주로 설계자의 판단에 의존한다.

⑤ 학습률(Learning Rate)

인공신경망의 학습 완급을 조정하는 상수인데 보통 0에서 1사이의 값을 취한다. 학습률을 크게 잡으면 학습속도가 빨라지지만 최소점을 지나칠 우려가 있다. 한편 학습률을 작게 잡으면 최소점에 조심스럽게 접근하기는 하지만 학습속도가 느려진다. 일반적으로, 학습이 진행되어 가면서 학습률의 값을 줄여 나간다.

⑥ 모멘텀(Momentum)

학습률의 크기에 따라 나타나는 양극단 현상을 완화시켜주기 위하여 도입된 개념이다. 즉, 전 단계의 연결강도의 일부를 현 단계의 연결강도에 반영함으로써 작은 학습률로 빠른 학습속도를 얻고자하는 것이다. 일반적으로, 학습이 진행되어가면서 모멘텀의 값을 줄여 나간다.

⑦ 연결정도(Connectivity)

층 사이의 처리요소들이 어느 정도 연결되었는가를 나타내는데, 층 사이의 처리요소들이 항상 모두가 연결될(fully connected) 필요는 없다. 층 사이의 모든 처리요소들을 연결해서 학습을 하더라도 학습이 끝난 후 연결정도를 관찰해보면 층 사이의 특정 처리요소간 연결은 불필요한 경우가 발생한다.

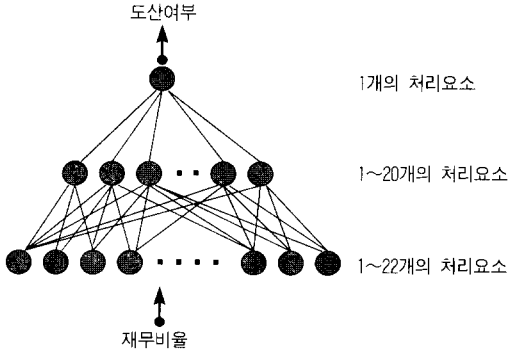
이와 같이 인공신경망은 그 설계에 있어서 고려하여야 할 요소가 많음에도 불구하고, 아직 정립된 방법론이 없는 실정이다. 그러므로 주로 설계자의 경험에 의존하거나, 은닉층의 처리요소의 수를 변화시키면서 테스트 적중률을 재고시키는 방향으로 구조를 개선시키는 방법이 사용되고 있다. 본 연구에서는 이에 유전적 알고리즘을 이용하여 인공신경망 구조의 설계 방법을 개선시키고자 하는데, 위에서 열거한 모든 고려요소들을 다루기에는 한계가 있으므로, 토폴로지에 초점을 맞추어 연구를 수행하고자 한다.

2.2 도산 예측을 위한 인공신경망

본 연구에서 유전적 알고리즘을 이용하여 설계하고자 하는 인공신경망은 [그림 2-2]와 같은 구조를 갖고 있다.

입력변수용 자료는 기업 도산예측시 고려할 수 있는 <표 2-1>에 제시된 22개의 재무비율을 이용하였는데, 이 재무비율들 중에서 어떠한 것들이 입력변수로 선정되어서 실제로 인공신경망의 학습 및 테스트용으로 사용되느냐에 따라 입력층의 처

리요소의 개수는 달라지게 된다.



[그림 2-2] 도산 예측을 위한 인공신경망의 구조

〈표 2-1〉 입력 변수용 재무비율

구분	재무비율
안정성 비율	(1)유동비율, (2)당좌비율, (3)고정비율, (4)고성장기적합율, (5)부채비율, (6)자기자본비율, (7)매출채권대매입채무비율
수익성 비율	(8)총자본수익율, (9)총자본경상이익율, (10)매출액순이익율, (11)매출액경상이익율
활동성 비율	(12)총자본회전율, (13)자기자본회전율, (14)경영자본회전율, (15)매출채권회전율, (16)매입채무회전율
생산성 비율	(17)부가가치율, (18)노동생산성, (19)총자본투자효율
성장성 비율	(20)총자본증가율, (21)매출액증가율, (22)유형고정자산증가율

은닉층의 처리요소의 개수도 입력변수의 개수와 그 수의 몇 %를 은닉층의 처리요소 개수로 할 것이냐에 따라 변하는데 1과 20 사이의 수로 결정된다. 출력층의 처리요소의 수는 도산이나 비도산이나를 나타내는 한 개의 처리요소를 가지고 있다. 서론에서 언급하였듯이 입력층과 은닉층의 처리요소의 개수는 유전적 알고리즘에 의하여 정해진다. 학습된 인공신경망의 성능은 테스트용 자료에 대한 예측 적중률로 판단하게 되는데, 테스트 적중률은 아래와 같은 식에 의하여 계산된다.

테스트 적중률 =

$$\frac{\text{도산기업은 도산으로 예측한 경우수} + \text{비도산기업은 비도산으로 예측한 경우수}}{\text{테스트용 기업 수}}$$

> 100

### 3. 유전적 알고리즘

유전적 알고리즘(Genetic Algorithm: GA)은 자연의 법칙인 '적자생존의 원리'에 근거를 두고 있다[Goldberg, 1989]. 즉 환경에 맞추어 생물들이 진화해 가는 과정을 알고리즘화한 것으로서 생물학에서 사용하는 용어를 그대로 사용하고 있다. 후보해(candidate solution)들의 집합을 개체군(population)이라고 하며, 이 후보해들은 염색체(chromosome)라고 불린다. 염색체는 유전자(gene)로 구성되고, 각 유전자들은 값을 갖는데, 이 값을 대립형질(allele)이라고 한다. 실제 문제영역에서 하나의 후보해가  $x_1, x_2, x_3$ 로 구성되어 있다고 하면, 유전적 알고리즘에서는 이것들의 값을 하나의 염색체로 표현한다. 그러므로 하나의 염색체는  $x_1, x_2$  및  $x_3$ 의 값을 표현하는 부분으로 구성되어 있는데, 각  $x_i$ 의 값에 해당하는 염색체의 부분을 유전자형(genotype)이라고 하고, 그 유전자형에 대응되는 실제 문제영역에서의 값, 즉  $x_i$ 의 값을 표현형(phenotype)이라고 한다. 하나의 염색체는 그에 따른  $x_i$ 의 값에 의하여 실제문제에서의 적응함수 값(fitness value)을 갖는다.

유전적 알고리즘이 해를 찾는 문제에 있어서 우수한 잠재력을 보이는 것은 재생산(reproduction), 교배(crossover), 그리고 돌연변이(mutation)라는 세 개의 유전자 조작자들 때문이다. 재생산 조작자는 교배와 돌연변이가 일어나기 전에 전 세대의 염색체들을 이용하여 다음 세대(generation)의 염색체들을 만드는 과정을 말하는데, 그 방법에 따라 세대적 재생산(Generational Reproduction : GR)과 안정상태 재생산(Steady-State Reproduction : SSR)으로 구분된다[Davis, 1991]. GR은 전 세대의 염색체들을 모두 다음 세대로 그대로 복사한 후, 그 염색체들에 교배와 돌연변이를 적용하여 새로운 염색체들로 모두 바꾸는 것을 말한다. 반면 SSR에서는 전 세대의 염색체들을 다음 세대로 일단 모두 복사한 후, 일정한 수의 염색체들을 제거한다. 그

다음, 남아있는 염색체들에 교배와 돌연변이를 적용하여, 제거한 수만큼의 새로운 염색체들을 생성시켜 삽입한다. 제거할 염색체의 수는 전체 개체군의 수보다 적으며, 보통 전 세대에서 일정 수준이하의 낮은 적응함수값을 갖거나 제일 오래된 염색체들을 제거한다.

SSR은 GR보다 효율적이다. 왜냐하면 초기화 과정만 제외하고는 한 세대에서 실제로 생성되는 염색체의 수가 GR보다 훨씬 적기 때문이다. 또한 전 세대에 있었던 우수한 적응함수값을 갖는 염색체들의 형질이 계속되는 세대에서도 유전될 확률이 높아지기 때문이다. GR에서는 전 세대의 염색체들을 다음 세대의 염색체들로 모두 대체해 버리므로 전 세대에서 출현한 우수한 염색체의 형질이 다음 세대에 유전될 확률이 낮아질 수 있다. 더욱이 염색체의 길이나 개체군의 크기가 매우 클 경우에, GR 방법에 의하여 재생산을 할 때마다 개체군의 크기만큼 염색체들을 만들어 낸다는 것은 해를 찾는 데 있어서 비효율적이라고 할 수 있다[Vavak and Fogarty, 1996]. 이러한 면에서 SSR은 적은 수의 염색체만을 재생산하면서도 우수한 염색체들을 유전시킬 수 있기 때문에 GR보다 효율적이며 경제적이라고 할 수 있다.

전 세대에서 복사된 염색체는 돌연변이가 일어나기 전에 교배가 진행된다. 어떤 염색체가 교배에 선택될 것인가 안될 것인가는 교배 확률에 의하여 결정된다.  $j$ 번째 세대의 염색체  $i$ 가 교배에 선택될 확률  $P_{ij}$ 는 아래와 같다.

$$P_{ij} = \frac{F_{ij}}{\sum_{i \in U} F_{ij}}, \quad i \in U, \quad j \in G.$$

- $F_{ij}$  : 적응함수 값
- $U$  : 염색체 개체군, 즉 Population
- $G$  : 세대, 즉 Generation

즉, 적응함수 값이 클수록 그 염색체가 교배용으로 선택될 확률이 커지게 되는데, 적응함수 값이 크다는 것은 염색체의 형질이 우수하다는 것이므로 결국 우수한 형질의 염색체가 다음 세대로 유전되는

확률이 커진다는 것이다.

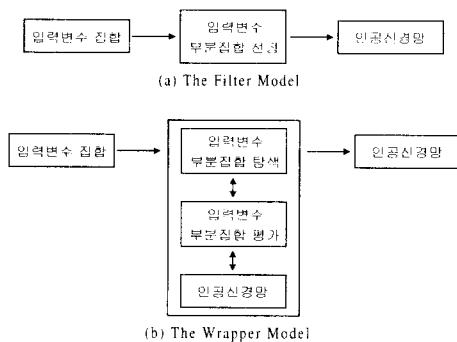
교배 확률에 의해서 두 염색체가 선택되면 교배할 임의의 위치를 선정하고, 이 위치 이후의 부분을 서로 교환한다. 교배가 일어날 위치  $k$ 는 난수에 의해 선정되는데, 교배가 일어날 두 염색체에서 이 위치는 같다. 교배의 위치  $k$ 는 두개 이상일 수도 있다. 교배 후에는 돌연변이 확률에 의해서 돌연변이가 일어난다. 돌연변이는, 예를 들어 유전인자들이 '0'과 '1'로 표현되어 있다면, '0'의 유전인자는 '1'로, '1'의 유전인자는 '0'으로 바꾸는 것이다. 돌연변이가 필요한 이유는 재생산이나 교배 조작자만으로는 생겨날 수 없는 잠재적인 해를 놓치지 않기 위함이다. 돌연변이가 일어날 확률은 아주 작게 정하는데 보통 0.01에서 0.08사이의 값, 혹은 더 작은 값을 사용하기도 한다.

#### 4. 인공신경망의 구조설계에 관한 기존의 연구

인공신경망의 설계시에 고려하여야 할 사항은 제2절에서 기술한 바와 같이 매우 다양하다. 그러나 이 요소들을 전부 고려할 수는 없으므로, 기존의 연구들을 살펴보면 구조설계 면에서는 주로 토폴로지나 연결정도 측면에 초점을 맞추었으며, 학습 방법의 면에서는 다른 기법을 이용하여 백프로퍼게이션의 단점을 극복하고자 하였다. 본 연구에서는 토폴로지의 측면에서 입력변수의 선정과 은닉층 처리요소의 수를 유전적 알고리즘에 의하여 결정하고자 한다. 입력변수를 선정하는 모델은 크게 Filter Model과 Wrapper Model로 구분할 수 있다. John *et al.*[1994]은 귀납적 학습알고리즘의 입력변수 선정방법을 연구한 그들의 논문에서 이 두 선정 모델의 차이를 비교하였는데, 인공신경망에 대상으로 하는 본 연구에서도 그들의 개념을 [그림 4-1]과 같이 무리 없이 적용할 수 있다.

두 모델의 구분은 사용될 입력변수의 선정작업이 학습알고리즘과 독립적으로 수행되느냐 아니냐에 달려 있다. 만일 입력변수의 선정이 [그림 4-1]

의 (a)와 같이 우리가 사용할 학습알고리즘인 인공 신경망과는 무관하게 결정된다면, 그 선정방법은 Filter Model이다. 하지만, [그림 4-1]의 (b)와 같이 입력변수의 선정과정에서, 사용할 학습알고리즘의 평가를 받으면서 선정된 입력변수의 부분집합이 계속 수정된다면, 이는 Wrapper Model이다. Filter Model이 Wrapper Model에 비하여 빠른 시간 내에 입력변수를 선정해주므로 효율적인 면에서는 우수하지만, Wrapper Model은 우리가 사용할 학습알고리즘의 특성을 반영한 입력변수를 선정해주므로 결과적인 성능은 Wrapper Model로 선정된 입력변수인 경우가 더 우수하다. 본 연구에서 제시하는 연구 모델도 Wrapper Model의 범주에 속하는데 이에 대해서는 제5절에서 자세히 기술한다. 본 절에서는 인공신경망 구조설계의 기존 연구들을 소개하며 본 연구와의 차이점도 더불어 언급하고자 한다.



[그림 4-1] 입력변수 선정모델 : Filter Model과 Wrapper Model

유전적 알고리즘을 인공신경망의 설계에 이용하는 연구는 주로 처리요소간의 연결정도 결정에 초점을 맞추어져 왔는데[Miller *et al.*, 1989; Balakrishnan and Honavar, 1995], Medsker[1995]는 인공신경망의 입력변수를 선정하는데 유전적 알고리즘을 이용하였다. 기본적인 연구 모델은 제5절에서 제시하는 본 연구의 연구 모델과 매우 흡사하지만 본 연구와는 두 가지 면에서 다르다. 첫째, 인공신경망의 구조를 결정함에 있어서 그들은 입력변수만을

유전적 알고리즘에 의하여 선정하였으며, 은닉층의 수는 1개이고 처리요소의 수는 4개로 고정시켜서 실험을 하였다. 둘째로, 본 연구에서는 제3절에서 소개한 유전적 알고리즘의 재생산 방법인 GR과 SSR을 모두 사용하여 그 결과를 비교하였는데, Medsker의 연구에서는 재생산 방법을 GR로만 사용하였다. Medsker는 원자력 발전소에 대한 7가지의 고장 상황 시나리오를 생성시켜 그의 연구 모델을 실험하였는데, 각 상황들에는 총 22개의 입력변수가 주어졌다. 그의 유전적 알고리즘에 의하여 입력변수는 평균 7개가 선정되었으며, 적중률은 평균 98%를 보였다. 하지만, 입력변수를 모두 사용하였을 때의 적중률이나 다른 선정방법의 적중률과의 비교는 제시되지 않았다.

Opitz and Shavlik[1996]은 하나의 인공신경망의 토폴로지 결정의 어려움을 극복하고자 여러 개의 인공신경망의 결과를 결합하여 하나의 결과를 산출하는 모형을 연구하였다. 이렇게 다양한 토폴로지를 가진 인공신경망들의 결합을 앙상블(Ensemble)이라고 부르는데, 앙상블내의 각 인공신경망은 두 가지의 조건을 만족하도록 구축되어야 한다. 첫째, 각 인공신경망은 혼란이 잘 되어있어 매우 정확한 결과를 산출하여야 한다. 둘째, 하나의 인공신경망은 다른 인공신경망이 오류를 많이 산출하는 입력자료에 대해서 정확하게 작동하여야 한다. 즉, 각각의 인공신경망은 입력자료 공간의 각각 다른 부분에 대해서 오류를 만들어 냄으로써 서로 오류 면에서 보완된다는 것이다. 이러한 다양한 토폴로지의 인공신경망들이 산출해낸 결과들을 단순 가중 평균하여 하나의 최종 결과를 산출해낸다. 그들은 앙상블에 추가될 새로운 인공신경망 토폴로지들을 생성해 내기 위하여 유전적 알고리즘인 REGENT[Opitz and Shavlik, 1994]를 개발하였다. REGENT는 인공신경망의 입력변수를 선정하기 위한 알고리즘이 아니다. REGENT의 유전적 알고리즘은 두 인공신경망의 은닉층에 대해서만 작용을 한다. 즉, 두 인공신경망의 은닉층의 처리요소에 대해서 교배와 돌연변이를 일으키는 것이

다. 그러므로 그들의 연구에서 유전적 알고리즘의 개체군에 있는 모든 인공신경망들은 동일한 개수의 입력변수를 갖고 있는 것이다. 그들은 REGENT 알고리즘으로 생성된 인공신경망들의 양상분을 인간 DNA상의 특정 유전자의 위치를 확인하는 연구에 적용하였다. 양상분의 성능이 개별적으로 사용된 인공신경망의 성능보다 우수하였는데, 적중률이 평균 약 2% 포인트 증가하였다.

Back *et al.*[1996]은 도산예측을 위한 인공신경망의 구조를 유전적 알고리즘을 사용하여 결정하였다. 그들의 연구가 적용 도메인에 있어서는 우리의 연구와 흡사하지만, 본 연구와는 두 가지 면에서 다르다. 첫째, 인공신경망의 구조를 결정함에 있어서 그들은 입력변수만을 총 31개의 입력변수 집합에서 유전적 알고리즘에 의하여 선정하였고, 은닉층의 처리요소의 수는 각각의 인공신경망마다 실험에 의하여 가장 좋은 적중률을 주는 개수로 결정하였다. 그 결과 은닉층의 수는 1개였으며, 은닉층의 처리요소 수는 7~12개 사이에서 결정되었다. 둘째로, 유전적 알고리즘의 조작자 면에서 재생산 방법을 GR로만 사용하였으며, 교배가 일어나는 위치의 수를 본 연구에서는 하나로 했는데 반해서 Back *et al.*은 두 군데 이상으로 하였다. 그들은 유전적 알고리즘에 의하여 입력변수가 선정된 인공신경망의 성능을 31개의 입력변수 모두를 사용하였을 때의 인공신경망의 성능과 비교하지는 않았다. 다만 Discriminant Analysis나 Logit Analysis를 입력변수의 선정기법으로 사용하여 구축한 인공신경망의 성능과 비교하였다. 그 결과 유전적 알고리즘을 사용하여 구축한 인공신경망이 적게는 2% 포인트로부터 많게는 16% 포인트의 적중률 증가를 보였다.

Yang and Honavar[1997]도 유전적 알고리즘을 사용하여 인공신경망의 입력변수를 선정하였다. 유전적 알고리즘으로 적합한 입력변수를 찾아내기 위해서는 수많은 다른 구조의 인공신경망을 훈련시켜야 하는데, 이 때에 전통적인 백프로피게이션 학습 알고리즘을 사용하면 효율적인 면에서 문제

가 생긴다고 그들은 주장한다. 그러므로 그들은 그들이 개발한 새로운 학습 알고리즘인 DistAI를 사용하였다[Yang *et al.*, 1997]. DistAI는 은닉층의 처리요소를 하나씩 추가시키면서 인공신경망의 적중률을 높여 가는 학습 알고리즘이다. 그러므로 그들의 연구에서는 유전적 알고리즘에 의하여 입력변수들만이 선정되며, 은닉층의 처리요소 수는 학습과정에서 결정되는 것이다. 그들은 유전적 알고리즘과 DistAI에 의하여 생성된 인공신경망을 공개된 기계학습용 자료인 UCI Data에 적용하였다[Blake *et al.*, 1997]. 그들이 사용한 UCI Data는 총 25종류인데 입력변수의 수나 자료집합의 크기 면에서 매우 다양하다. 그들이 생성한 인공신경망의 성능을 입력변수를 모두 사용하였을 때의 인공신경망의 성능과 평균적인 면에서 비교하여 보면, 유전적 알고리즘에 의하여 입력변수는 원래 자료의 입력변수 수의 1/2정도가 선정되었으며, 적중률은 82.8%에서 90.6%로 향상되었다. 하지만, 은닉층의 처리요소 수는 입력변수를 모두 사용하였을 때보다 약 40%가 증가하였다. 그들은 UCI Data의 일 부분에 대해서는 Decision Tree 알고리즘의 성능과 비교하였는데, 그들의 알고리즘이 Decision Tree 알고리즘보다 평균 2배정도 많은 입력변수를 사용하였지만, 평균 적중률은 Decision Tree의 81.2%보다 높은 93.4%를 기록하였다.

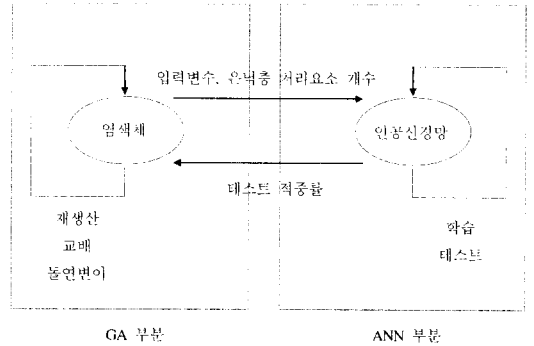
Setiono and Liu[1997]는 본 연구와는 달리 인공신경망의 입력변수 선정에 있어서 유전적 알고리즘을 사용하지 않고, 인공신경망 자체에서 얻은 정보를 이용하였다. 즉, 인공신경망의 처리요소간의 연결강도를 변경하면서 관련성이 적은 입력변수(Irrelevant Attribute)를 제거하는 알고리즘을 개발하였다. 먼저 모든 입력변수를 사용하여 은닉층이 하나인 3계층 인공신경망을 훈련시킨 후 적중률을 구한다. 이 값을  $P_R$ 이라 하자. 그 다음, 각 입력변수  $A_i$ 에 대해서,  $A_i$ 에 관련된 모든 연결강도를 0으로 놓은 후 인공신경망을 훈련시켜서 적중률을 구한다. 이 값을  $P_i$ 라 하자. 이렇게 하여

모든 입력변수들에 대하여 해당 적중률들을 구한 후에  $P_R - P_i$ , 즉 적중률 하락이 가장 작은 입력 변수를 제거한다. 이러한 과정을 남아있는 입력변수들에 대하여 반복하면서 적중률 하락이 미리 정한 한계를 넘어서기 전까지 계속하는 것이다. Setiono and Liu는 그들의 알고리즘을 UCI Data를 포함한 여러 종류의 데이터에 적용하였다. 입력변수의 수는 그들의 알고리즘에 의하여 결정되었지만, 은닉층 처리요소의 수는 어떤 데이터던 간에 12개로 고정시켰다. 그들이 사용한 4개의 UCI Data에 대한 결과를 보면, 입력변수의 수가 평균 20%의 수준으로 감소하였는데, 그 중 Sonar Target Dataset의 경우에는 특이하게도 입력변수의 수가 원래의 60개에서 4개 정도로 대폭 감소되었는데도 적중률은 60개의 입력변수 전부를 사용하였을 때의 92.3%보다 약간 증가한 93.8%를 보였다. 평균적으로, 적중률은 모든 입력변수를 전부 사용하였을 때보다 약 2% 포인트 증가하였다. Yang and Honavar[1997]가 동일한 데이터에 대하여 수행한 결과와 비교하여 보면, Setiono and Liu의 알고리즘이 더 적은 수의 입력변수를 선정하고 있지만, 적중률에서는 Yang and Honavar의 결과가 5% 포인트 정도 높았다.

### 5. 유전적 알고리즘으로 설계한 인공신경망 : ANN\_D\_GA

본 연구에서 고안한 연구 모델인 ANN\_D\_GA (Artificial Neural Network Designed using Genetic Algorithm)는 [그림 5-1]에서 보는 바와 같이 크게 두 부분, 즉 GA부분과 ANN부분으로 구성되어 있다. GA부분에서는 유전적 알고리즘을 이용하여 인공신경망의 입력변수들을 선정하고 은닉층 처리요소 개수를 결정한다. ANN부분에서는 GA부분의 결과에 따라 인공신경망을 구성하여 학습한 후 테스트를 하게 된다. 특히 GA부분에서는 재생산 방법의 차이를 보기 위해 GR과 SSR 방법

을 이용하여 두 종류의 실험을 한 후 그 결과를 비교하였다.



[그림 5-1] ANN\_D\_GA의 GA부분과 ANN부분

#### 5.1 GA 부분

인공신경망의 입력변수로는 <표 2-1>에서 보는 바와 같은 기업 도산예측시 고려할 수 있는 22개의 재무비율들 중에서 선정하였는데, 각 재무비율의 선정여부는 0 또는 1의 이진수(bit)로 표현하였다. 그러므로 재무비율의 선정여부를 표현하기 위해서는 총 22 bits가 필요하다. <표 2-1>의 각 재무비율에 붙어있는 번호는 검색체에서의 자리 수를 나타낸다. 즉 어떤 검색체에서 10번째 유전인자의 값이 0이면 매출액순이익율이 입력변수로 선정되지 않았음을, 1이면 매출액순이익율이 입력변수로 선정되었음을 표현한다.

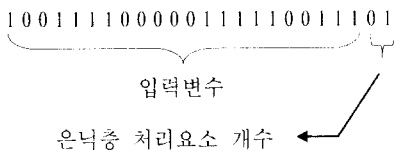
Salchenberger *et al.*[1992]은 은닉층 처리요소 개수로 입력변수 개수의 75%를 사용하는 것이 테스트 적중률을 향상시킨다는 연구를 발표하였다. 본 연구에서는, 은닉층 처리요소 개수가 선정된 입력변수 개수의 25%, 50%, 75% 또는 90%가 되도록 실험을 설계하였는데, 이러한 네 가지 경우를 표현하기 위하여 2 bits가 필요하다. <표 5-1>은 은닉층의 처리요소 개수를 표현하기 위한 유전자 형과 이에 대응하는 표현형을 보여주고 있다. 실제로 인공신경망에서의 은닉층 처리요소 개수는 선정된 입력변수의 개수에 표현형을 곱한 수를 반올림하여 결정하였다.



<표 5-1> 은닉층 처리요소 개수를 위한 유전자형과 표현형

유전자형	표현형
00	0.25
01	0.50
10	0.75
11	0.90

즉, 하나의 염색체는 재부비율의 선정여부를 표현하는 22 bits와 은닉층 처리요소 개수를 결정하기 위한 2 bits인 총 24 bits로 구성되어 있다. [그림 5-2]는 염색체의 한 예를 보여주고 있는데, 이 염색체의 정보를 해석하여 만들어진 인공신경망은 13개의 입력층 처리요소, 7개( $13 \times 0.5 = 6.5$ 의 반올림)의 은닉층 처리요소와 1개의 출력층 처리요소로 구성된다. 이 때에 선정된 입력변수들은 <표 2-1>에서 1, 4, 5, 6, 7, 13, 14, 15, 16, 17, 20, 21, 22번의 재부비율들이다.



[그림 5-2] 염색체의 예

이러한 염색체 20개가 모여서 한 세대의 개체군을 구성한다. 제1세대의 염색체들은 난수에 의하여 생성시켰고, 제2세대부터는 재생산, 교배, 그리고 돌연변이 등의 유전자 조작을 통해서 생성시켰다. 적응함수값은 ANN부분에서 얻어지게 되는데 인공신경망의 테스트 적중률을 그대로 사용하였다. 본 연구에서 사용한 교배확률은 0.65이고 돌연변이 확률은 0.08이다. GR을 이용해서 실험을 했을 경우에는 각 세대마다 20개의 염색체가 생성된다. 하지만, SSR의 경우에는 전 세대에서 복사된 염색체들 중에서 적응함수값이 가장 작은 순서로 6개의 염색체를 제거하고, 나머지 14개의 염색체들을 바탕

으로 유전자 조작을 통해서 6개의 염색체를 만들어 낸다. 그러므로, SSR의 경우에는 제1세대에서는 GR과 마찬가지로 20개의 염색체가 생성되지만, 제2세대 이후에는 6개의 염색체가 생성된다.

### 5.2 ANN 부분

GA부분에서 인공신경망의 입력변수, 은닉층 처리요소 개수 등을 위한 염색체가 생성되면 ANN부분에서는 이를 바탕으로 인공신경망을 구성한다. 구성된 인공신경망은 GA부분에서 선정된 입력변수 개수만큼의 입력 처리요소, <표 5-1>의 표현형에 의하여 계산된 개수만큼의 은닉 처리요소, 그리고 한 개의 출력 처리요소를 가지고 있으며 백프로퍼게이션 방법에 의하여 학습된다. 하나의 인공신경망의 학습은 최대 10만번까지 하도록 하였는데 중간에 50번의 주기로 테스트 적중률을 측정하였다. 하지만 5만 번이 넘은 이후에 100번 이내에 테스트 적중률이 갱신되지 않으면 학습을 멈추도록 하였다.

GA부분과 ANN부분의 과정을 통합해서 ANN\_D\_GA를 하나의 프로시저로 나타내면 다음과 같다.

- ① 난수를 이용하여 제1세대의 염색체들을 생성한다.
- ② 이 염색체들을 이용하여 인공신경망을 구성하고 학습한 후 테스트 적중률을 구한다.
- ③ 미리 정해 놓은 세대수에 도달하였으면 해 찾기 과정을 끝내고, 그렇지 않으면 ④번 과정으로 진행한다.
- ④ 유전자 조작자들을 이용하여 다음 세대를 구성한다.
- ⑤ 세대수를 하나 증가시키고 ②번 과정으로 돌아간다.

## 6. ANN\_D\_GA의 성능 실험

ANN\_D\_GA의 성능을 실험하기 위한 자료는 한

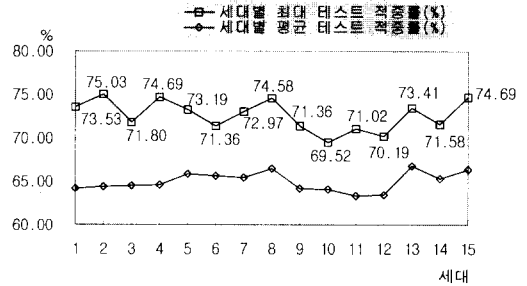
국신용보증기금으로부터 받은 1991년과 1993년 사이에 도산한 기업 61개와 비도산기업 59개를 사용하였다. 학습과 테스트를 위해서 자료들을 나누었는데 학습 자료는 도산기업 30개, 비도산기업 30개로 구성되어 있고 테스트 자료는 도산기업 31개, 비도산기업 29개로 구성되어 있다. 한 개의 자료는 <표 2-1>에 있는 22개의 재무비율과 도산 유무를 나타내는 데이터로 구성되어 있다.

각 세대의 염색체의 수는 20으로 정하였는데, 재생산 방법을 GR로 사용했을 때에는 세대수를 15로 하였다. 그러므로, 총 300개(= 20×15)의 인공신경망이 만들어졌다. 재생산 방법을 SSR로 사용했을 때에는, 새로 생성되는 인공신경망의 수가 적어서 인공신경망을 학습하고 테스트하는 계산과정의 부담이 적으므로 좀더 많은 세대인 25세대까지 진화시켰다. 그러므로 SSR을 사용할 때에는 총 164개(= 20+6×24)의 인공신경망이 만들어졌다. 이 인공신경망들의 테스트 적중률의 범위는 <표 6-1>과 같이 나타났다.

<표 6-1>에서 보듯이 개별적이나 세대별 평균 면에서 볼 때에 SSR이 GR에 비하여 다소 편차가 크지만, 전반적으로 SSR이 높은 테스트 적중률을 보이고 있다. 즉, SSR을 사용한 인공신경망의 테스트 적중률이 빠른 속도로 증가함을 알 수 있다.

[그림 6-1]은 GR을 적용했을 때의 세대별 최대 테스트 적중률과 평균 테스트 적중률의 추이를 나타내고 있다. 15개 세대의 총 300개의 인공신경망 중에서 테스트 적중률이 가장 좋은 것은 2번째 세대의 9번째 염색체로서 75.03%의 수치를 기록하였다. 이 최고값은 그 후에 갱신되지 않았다. 평균 테스트 적중률은 증가 추세를 보이다가 8번째 세대

를 지나서 하향 곡선을 그렸다.



[그림 6-1] GR을 사용한 ANN\_D\_GA의 테스트 적중률

생성되는 인공신경망들 중에서 입력변수 개수의 75%를 은닉층 처리요소 개수로 한 것이 70%를 넘어서서, Salchenberger *et al.*[1992]의 연구 결과를 뒷받침해 주었다. [그림 6-2]는 SSR을 적용했을 때의 실험 결과인데, GR을 사용했을 때와는 다른 결과를 볼 수 있다. 우선 세대별 테스트 적중률의 평균을 보면 대체적으로 증가하는 경향을 나타내고 있다. 더욱이 제25번째 세대의 평균 테스트 적중률 74.81%는 제1세대의 최대 테스트 적중률 74.69%를 넘어서고 있다. 최대 테스트 적중률의 추이는 19번째 세대까지는 큰 변동이 없었으나 20번째 세대를 지나면서 17000번 정도의 학습으로 78.03%의 테스트 결과를 보여 주었다. SSR이 GR보다 많은 세대를 진화하였음에도 새로 만들어진 인공신경망의 수는 적으며, 또한 실험 결과도 SSR이 GR보다 좋은 것을 고려해 보면 SSR이 GR보다 경제적인 것을 알 수 있다. 이것은 제1세대만 제외하고 나머지 세대에서 만들어지는 염색체의 수에 있어서 SSR의 경우가 GR의 경우보다 적기 때문이다.

<표 6-1> 재생산 방법별 ANN\_D\_GA의 테스트 적중률 범위

	개별적 인공신경망		세대별 평균		세대별 최대	
	최소	최대	최소	최대	최소	최대
GR	52.28	75.03	63.31	66.73	69.52	75.03
SSR	55.34	78.03	66.23	74.81	74.69	78.03

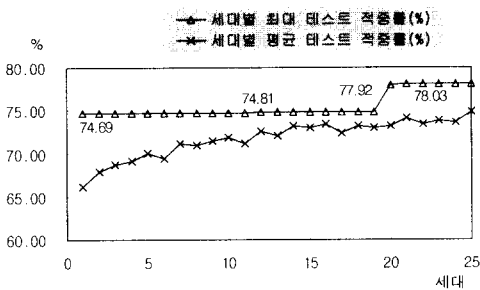
단위 : %

<표 6-2> SSR을 사용한 ANN\_D\_GA의 세대별 구조와 성능

세대번호	입력변수 개수	은닉층 처리요소 개수*	최대 테스트 적중률	선정된 재무비율 번호
1~11	8	6	74.69%	6 8 9 13 14 19 21 22
12	9	8	74.81%	2 3 5 8 13 14 19 21 22
13	10	8	74.81%	2 3 4 8 9 13 14 19 21 22
14	12	6	74.81%	2 3 4 5 8 9 10 11 13 14 19 21
15	9	8	74.81%	2 3 5 8 13 14 19 21 22
16~18	8	7	74.81%	1 5 8 13 14 19 21 22
19	12	9	74.81%	1 2 3 5 8 9 10 11 13 14 19 21
20	15	8	77.92%	2 3 4 5 6 8 9 10 11 12 13 14 17 20 21
21~25	8	6	78.03%	6 8 9 13 17 19 21 22

\* 최대의 테스트 적중률을 보인 인공신경망의 은닉층 처리요소 개수

<표 6-2>는 SSR을 사용한 ANN\_D\_GA의 세대별 입력변수의 개수, 그에 따른 은닉층의 처리요소 개수, 테스트 적중률 중 최대값, 그리고 선정된 재무비율의 번호를 보여 주고 있다. 가장 성능이 우수한 인공신경망은 제 21 세대 이후에 얻어지고 있다.



[그림 6-2] SSR을 사용한 ANN\_D\_GA의 테스트 적중률

<표 6-2>에서 보듯이, 8개가 선정된 제1세대부터 11세대까지의 입력변수 구성과 가장 우수한 성능을 보인 제21세대 이후의 입력변수 구성을 비교해 보면, 단지 14번째 재무비율인 경영자본회전율이 17번째 재무비율인 부가가치율로 바뀐 것뿐이

다. 또한 제20세대에서는 15개의 입력변수가 선정되어 가장 복잡한 인공신경망이 구성되었지만, 그 성능이 가장 우수하지는 않았다. 그러므로, 입력변수의 수가 많다고 해서 성능이 우수한 인공신경망이 구성되는 것이 아니며, 동일한 수의 입력변수를 사용하더라도 선정된 재무비율에 따라 인공신경망의 성능에 차이를 보이고 있다.

입력변수 선정에 관한 효과를 비교하기 위하여 본 연구와 과거에 발표된 인공신경망들 간의 테스트 적중률을 비교하여 <표 6-3>에 제시하였다.

먼저, 유전적 알고리즘을 사용하지 않고 <표 2-1>에 있는 22개의 입력변수를 모두 선정하고 은닉층 처리요소 개수는 17개 즉  $22 \times 0.75 = 16.5$ 를 반올림하여 설정한 다음 인공신경망을 테스트하였다. 결과는 64.57%의 적중률을 보였다. 이재식과 한재홍[1996]의 연구에서는 본 연구에서 사용한 22개의 재무비율을 모두 입력변수로 사용하고 은닉층의 처리요소 수를 5개로 설정하여 20만 번을 학습한 결과 70%의 적중률을 기록하였다. 그들의 또 다른 연구[이재식과 한재홍, 1995]에서는 본 연구에서 사용한 22개의 재무비율 중에서 생산성 관련비율

<표 6-3> 인공신경망의 연구 모형별 테스트 적중률

구분	연구 모형	본 연구			이재식과 한재홍의 연구	
		GR	SSR	GA 사용안함	1995년	1996년
입력변수의 수		8	8	22	22	22
은닉층 처리요소 수		6	6	17	10 또는 20	5
테스트 적중률		75.03%	78.03%	64.57%	64.0%	70.0%

을 제외하고 그 대신 수익성과 안정성에 관련된 재무비율의 개수를 늘여서 총 22개의 재무비율을 사용하여 도산예측을 위한 인공신경망을 개발하였다. 은닉층의 처리요소 수와 학습회수에 따라 약간의 차이를 보였지만, 이들 인공신경망의 테스트 적중률은 64% 정도를 보였다.

## 7. 결 론

인공신경망이 다양한 분야에서 활용되어 잠재력을 인정받고는 있지만 인공신경망의 설계에 있어 뚜렷한 기준이 없는 것은 단점이 아닐 수 없다. 본 연구에서는 인공신경망의 입력변수의 선정과 은닉층 처리요소 개수의 설정을 유전적 알고리즘에 의하여 수행함으로써 성능이 개선된 인공신경망을 설계하였다. 본 연구에서 고려한 22개의 입력변수를 모두 사용한 인공신경망의 경우에는 테스트 적중률이 70%를 넘지 못하였으나, 유전적 알고리즘을 이용하여 입력변수를 선정한 경우에는 최고 78.03%의 테스트 적중률을 기록하였다. 이것은 입력변수를 모두 사용하여 예측을 하는 것보다는 설명력이 있는 입력변수만을 선정하여 예측을 수행하는 것이 좀더 향상된 예측력을 나타낸다는 것을 의미한다. 또한 인공신경망의 테스트 적중률은 은닉층의 처리요소 개수와 무관하지 않는데 최대 테스트 적중률을 보이는 인공신경망들의 은닉층 처리요소 개수를 관찰해 보면 선정된 입력변수 개수의 75%를 사용한 경우가 가장 많았다.

본 연구의 한계점을 지적하면 다음과 같다. 첫째로는 인공신경망의 은닉층의 수와 은닉층의 처리요소 개수를 다양하게 고려하지 못했다는 점이다. 본 연구에서는 은닉층의 수를 한 개로 하였으며, 그 처리요소의 개수도 입력변수 개수의 25%, 50%, 75%, 또는 90%로 설정하였다. 그러나 은닉층은 한 개 이상일 수 있으며, 그 처리요소의 개수도 비록 입력변수 개수의 75%일 때에 가장 효과가 있다고는 하지만 입력변수의 개수보다 많게 설정될 수도 있는 것이다. 이러한 점에서 본 연구는 인공신경망

의 다양한 토폴로지를 수용하지 못한 한계를 가지고 있다. 두 번째로 지적할 것은 유전적 알고리즘의 설계에 관한 것인데, 적응함수, 재생산 방법, 개체군의 크기 등에서 아쉬운 점이 있다. 본 연구에서는 인공신경망의 테스트 적중률을 그대로 적응함수값으로 사용하였는데 이러한 계산 방법은 그해가 도출된 개체군의 특성을 반영하지 못한다는 단점을 안고 있다. 이러한 단점을 극복하기 위해서는 세대수, 선정된 입력변수의 개수 등을 고려해서 적응함수를 도출하여, 동일한 테스트 적중률에 대해서도 세대수나 선정된 입력변수의 개수에 따라 다른 적응함수값이 계산되어야 할 것이다. 본 연구에서는 SSR에서 제거하는 염색체의 수를 6개로 하였지만 어떠한 기준을 두고 한 것은 아니었다. SSR은 유전자 조작을 통해서 새롭게 만드는 염색체의 수가 GR에 비해서 상대적으로 적어서 효율적이라는 장점이 있지만 새로운 우성인자가 나타날 기회가 줄어든다는 단점을 갖고 있다. 제거하는 염색체의 수를 얼마로 할 것인가 즉, 새롭게 생성되는 염색체의 수를 얼마로 할 것인가에 대해서는 또 다른 추가적인 연구가 필요하다. 본 연구에서는 개체군의 크기를 20으로 하였다. 유전적 알고리즘의 설계에 정해진 표준은 없지만, 하나의 염색체의 크기가 24bits인데, 개체군의 크기가 이보다 작았다는 것은 다양한 염색체가 생성될 기회가 제한되지 않았나 생각된다. 마지막으로 지적할 것은 기존 연구들에서 제시된 입력변수 선정방법과의 비교를 하지 못했다는 것이다. 제 4절에서 기존 연구의 모델과 본 연구의 모델의 차이점들을 기술하는데 그쳤으며, 기존의 방법을 구현하여 본 연구에서 사용한 자료에 적용하지는 못하였다.

## 참 고 문 헌

- [1] 이재식, 한재홍, "인공신경망을 이용한 중소기업 도산 예측에 있어서의 비재무정보의 유용성 검증", 「한국전문가시스템학회지」, 창간호 (1995), pp.123-134.

- [2] 이재식, 한재홍, "사례기반추론을 이용한 중소기업 도산 예측에 있어서의 비재무정보의 활용", 「한국전문가시스템학회 '96 추계학술대회 논문집」(1996), pp.243-252.
- [3] 한인구, 권영식, 이건창, 주성도, "지능형 기업 신용평가 시스템의 개발: NICE-AI", 「한국경영정보학회 '94 추계학술대회 논문집」, (1994), pp.229-252.
- [4] Balakrishnan, K. and V. Honavar, *Evolutionary Design of Neural Architectures*, Techn. Rep. No. 95-01, AI Research Group, Dept. of Computer Science, Iowa State Univ. 1995.
- [5] Bansal, A., R. J. Kauffman, R. M. Mark and E. Peters, "Financial Risk and Financial Risk Management Technology," *Information & Management*, Vol.24(1993), pp.267-281.
- [6] Back, B., T. Laitinen and K. Sere, "Neural Networks and Genetic Algorithms for Bankruptcy Predictions," *Proc. the 3rd World Congress on Expert Systems*, Seoul, Korea, Vol.1(1996), pp.123-130.
- [7] Blake, C., E. Keogh and C. J. Merz, *UCI Repository of Machine Learning Databases*, Dept. of Information and Computer Science, UC at Irvine, CA, [<http://www.ics.uci.edu/~mlearn/MLRepository.html>], 1997.
- [8] Davis, L., *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, 1991.
- [9] Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Pub. Co., Inc., 1989.
- [10] John, G. H., R. Kohavi and K. Pflieger, "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.
- [11] Kamijo, K. and T. Tanigawa, "Stock Price Pattern Recognition: A Recurrent Neural Network Approach," *Proc. Int'l Joint Conf. on Neural Networks*, Vol.1(1990), pp.215-222.
- [12] Medsker, L. R., "Genetic Algorithms to Select Inputs to Neural Networks," *Hybrid Intelligent Systems*, Kluwer Academic Publishers(1995), pp.162-169.
- [13] Miller, G. F., P. M. Todd and S. U. Hedge, "Designing Neural Networks using Genetic Algorithms," *Proc. the 3rd Int'l. Conf. on Genetic Algorithms*(1989), pp.379-384.
- [14] Nelson, M. M. and W. T. Illingworth, *A Practical Guide to Neural Nets*, Addison-Wesley Pub. Co., Inc., 1991.
- [15] Odom, M. D. and R. Sharda, "A Neural Network Model for Bankruptcy Prediction," *Proc. Int'l Joint Conf. on Neural Networks*, Vol.2(1990), pp.163-168.
- [16] Opitz, D. W. and J. W. Shavlik, "Using Genetic Search to Refine Knowledge-based Neural Networks," *Proc. of the 11th Int'l Conf. on Machine Learning*, New Brunswick, NJ, Morgan Kaufmann(1994), pp.208-216.
- [17] Opitz, D. W. and J. W. Shavlik, "Generating Accurate and Diverse Members of a Neural-Network Ensemble," in *Advances in Neural Information Processing Systems 8*, Touretzky, D. S., M. C. Mozer and M. E. Hasselmo (eds.), MIT Press, Cambridge, MA, 1996.
- [18] Salchenberger, L. M., E. M. Cinar and N. A. Lash, "Neural Networks: A New Tool for Predicting Thrift Failures," *Decision Sciences*, Vol.23(1992), pp.899-915.
- [19] Setiono, R. and H. Liu, "Neural-network

- Feature Selector," *IEEE Trans. on Neural Networks*, Vol.8, No.3(1997), pp.654-662.
- [20] Vavak, F. and T. C. Fogarty, *Comparison of Steady State and Generational Genetic Algorithms for Use in Nonstationary Environments*, Techn. Rep., Dept. of Computer Studies and Mathematics, U. of West of England, [<http://www.btc.uwe.ac.uk/evol/techniques.html>], 1996.
- [21] Yang, J. and V. Honavar, "Feature Subset Selection using a Genetic Algorithm," *Proc of the Genetic Programming Conf., GP'97*, Stanford Univ., CA, [<http://www.cs.iastate.edu/~honavar/papers/expert.ps>], (1997), pp.380-385.
- [22] Yang, J., R. Parekh and V. Honavar, *DistAI : An Inter-pattern Distance-based Constructive Learning Algorithm*, Techn. Rep., ISU-CS-TR 97-05, Iowa State Univ., 1997.