

# 인공신경망을 이용한 기업도산예측

- IMF후 국내 상장회사를 중심으로 -

정유석\*, 이현수\*\*, 채영일\*\*, 서영호\*\*\*

\*금융감독원, \*\*경희사이버대학교 글로벌경영학과, \*\*\*경희대학교 경영학부

## A Neural Network Model for Bankruptcy Prediction

-Domestic KSE listed Bankrupted Companies after the foreign exchange crisis in 1997-

Yu-Seok Jeong\*, Hyun-Soo Lee\*\*, Young-Il Chae\*\*, Yung-Ho Suh\*\*\*

\*Financial Supervisory Service, \*\*Kyunghee Cyber Univ., \*\*\*Kyunghee Univ.

### Abstract

This paper is concerned with analysing the bankruptcy prediction power of three models: Multivariate Discriminant Analysis(MDA), Logit Analysis, Neural Network. The after-crisis bankrupted companies were limited to the research data and the listed companies belonging to manufacturing industry was limited to the research data so as to improve prediction accuracy and validity of the model. In order to assure meaningful bankruptcy prediction, training data and testing data were not extracted within the corresponding period. The result is that prediction accuracy of neural network model is more excellent than that of logit analysis and MDA model when considering that execution of testing data was followed by execution of training data.

## 1. 서론

97년 11월 IMF 구제금융(이하 'IMF') 이후 정부는 기업의 부채비율 축소 등 주채무계열의 재무구조개선을 기업구조조정 우선 과제중의 하나로 삼았다. 또한 정부는 2001년 1월 기업의 재무구조개선 뿐만 아니라 기업 부실에 대한 사전적 감시체제를 구축하고 부실확대를 조기에 차단하여 금융회사의 건전성 제고와 시장기능에 의한 상시적 기업구조조정의 활성화를 도모하기 위해 2단계 기업구조조정으로 기업신용위험 상시평가제도를 도입하였다.

이에 은행은 기업의 신용위험을 상시적으로 평가하기 위해 기업신용등급 평가 모형의 근간이 되는 기업도산 예측 연구에 관심이 높아졌다. 또한 기업의 도산을 어떻게 예측할 것인가, 또한 예측이 가능한 일인가에 대한 것이 대출을 하는 금융회사 뿐 만 아니라 차입을 해야 하는 기업입장에서도 큰 관심사로 떠오르게 되었다.

기업도산예측 연구가 학계에서 지속적으로 연구되어 왔음에도 불구하고, 업계인 은행이 상시평가제도를 도입함으로써, 관련분야에서도 도산예측 연구의 필요성이 다시 부각되고 있는 실정이다.

본 연구의 목적은 IMF후에 도산한 기업을 대상으로 다변량판별분석 모형, 확률모형(로지트분석모형) 그리고 인공신경망 모형을 개발하여 각 모형의 도산예측력을 비교하고 인공신경망 모형의 일반화 가능성을 높이는 데 있다. 연구표본을 IMF후에 도산한 기업으로 한정할 이유는 유사한 경제환경의 영향을 받은 표본을 대상으로

모형을 개발할 경우 도산 예측력을 개선시킬 수 있으며, 모형의 일반화 가능성을 높일 수 있기 때문이다.

본 연구는 도산예측 모형간의 예측력 비교 측면에서는 기존 연구(이진창, 1993; Back et al., 1996)와 유사하나 연구표본을 IMF후에 도산한 기업으로 하여 도산예측력을 향상시키고 모형의 일반화 가능성을 높이기 위해 상장회사 중 동일한 업종인 제조업종에 한정하여 모형을 개발한다는 측면에서 기존 연구와 차이가 있다고 할 수 있다. 또한, 보다 의미있는 연구를 위하여 학습용 표본과 검증용 표본을 동일한 기간에서 추출하지 않고 검증용 표본을 학습용 표본기간 이후의 기간에서 추출하여 도산예측의 타당성을 현재가 아닌 미래의 시점에서 검증함으로써, 개발한 모형이 미래의 환경변화에 적응력을 보이는지를 분석하고자 한다.

## 2. 우리나라 기업도산에 관한 이론적 고찰

### 2.1. 상장회사의 부도발생 현황

상장회사의 부도발생추이를 보면 <표 1>과 같다. 96년에는 부도발생업체가 6개이나 외환위기 발생연도인 97년 한해에만 48개사에 달하였으며 2002년 1월부터 10월까지 부도발생 업체 수는 12개이다. 이는 부도발생 업체의 수가 외환위기 직후보다 다소 줄어든 편이나 외환위기 전보다는 늘어났음을 알 수 있다.

<표 1> 상장회사의 부도발생수 추이

(단위: %, 개)

연도	비율 (A/B)	부도업체수(A)		총상장회사수 (B)
			제조업종	
1996	0.8	6	2	760
1997	6.2	48	40	776
1998	6.7	50	42	748
1999	1.1	8	4	725
2000	2.1	15	10	704
2001	3.9	27	24	689
2002	1.8	12	9	683
합계		166	131	-

자료원: 한국증권거래소, 한국상장사협의회

97년 초 한보, 삼미, 진로 등 대기업을 중심으로 부도가 발생하여 많은 하청 중소기업이 부도처리되었고 이는 외환위기와 결부되어 금융권의 붕괴를 초래하였다. 이로 인해 금융권은 자금지원을 상당히 보수적으로 운영하게 되었으며 유동성이 부족한 기업은 대기업을 막론하고 부도처리되었으며 98년의 부도발생 업체 수는 97년보다 증가한 50개사에 이르게 되었다.

98년 6월 정부는 채권금융회사 자율협약인 워크아웃제도를 도입하였으며 부도처리된 상장회사는 회사정리절차나 회의를 통해 기업갱생을 도모하였다. 그 결과 99년 상장회사 부도발생 업체 수는 8개사로 97·98년에 비해 대폭 감소하였다. 그러나 2000년 중반이후 대우그룹 경영부실에 따른 파급영향, 국내외 경기침체 등 대내외 여건 악화로 금융시장의 불안이 재연됨에 따라 정부는 2단계 금융 및 기업구조조정을 추진하였고 그 영향으로 2001년의 상장회사 부도발생 업체 수는 27개사로 99년이후 가장 높은 수치를 보였다. 2002년 10월 상장회사 부

도발생 업체 수는 11개로 많이 감소하였으나 우리나라 경제가 아직 완전히 회복되지 않았기 때문에 당분간 부도발생 업체 수는 증가할 것으로 추정된다.

한편 상장회사의 부도발생 업체 중 제조업이 차지하는 비중은 평균 77.4%로 상당한 부분을 차지하고 있다. 97·98년에는 경기불황에 따른 수익성 악화로, 특히 1차 및 조립금속 업체(제조업 중 섬유·의복·가죽 등)는 현금유출이 지속됨에도 불구하고 자산매각이나 설비투자 축소와 같은 구조조정을 제대로 추진하지 못하였다. 또한 투자활동에서의 현금유출이 계속 증가하여 필요자금을 외부차입 등에 과도하게 의존하게 된 상태에서 직접금융시장의 자금조달 사정이 급격히 악화되어 부도에 이르게 되었다.

그러나 2002년 상반기 제조업의 부채비율은 135.6%로 98년 303.0%에 비해 1/3 수준으로 감소하였다. 이렇게 제조업의 재무구조가 개선된 것은 구 대우그룹 관련기업들의 구조조정이 반영된 데다, 수익성이 크게 개선되고 우량 대기업을 중심으로 차입금을 적극 상환하였기 때문이라 할 수 있다.

## 2.2. IMF전·후 기업도산의 원인 비교

IMF전·후 기업도산의 원인을 연구자별로 열거하면 다음과 같다.

남명수(1998)의 연구에 의하면 IMF전 부도기업의 재무적 특징은 정상기업에 비해 상당히 많은 부채를 조달하였고 금융비용이나 영업상 이자지급규모가 크다는 것이다. 특히 지속적인 수익성 악화로 부채조달규모는 점점 상승하였고, 어느 정도의 단기지급능력을 보유하고 있었음에도 불구하고 이자지급능력 부족과 부채상환능력 부족으로 부도가 발생하였다는 것이다. 그러나 IMF후의 부도기업은 97년도에 급격히 수익성이 하락하여 단기 운영자금을 금융이나 사채 등의 타인자본을 이용한 것이 아니라 자기자본과 유동부채를 통해 조달한 것이 직접적인 부도원인으로 나타났다. 이에 IMF후 부도기업은 수익회복능력이 있었음에도 불구하고 단기간의 유동성 악화와 이로 인한 단기지급능력의 부족으로 부도가 발생한 것이라고 하였다.

결국 IMF전의 부도원인은 과거의 인일한 경영정책으로 시작된 현저한 재무상태의 악화와 환경변화에 적응할 수 있는 대응력의 부족인 반면 IMF후의 부도기업은 재무상태가 건전함에도 불구하고 단기자금조달에 실패함으로써 부도가 발생했음을 알 수 있다. 따라서 IMF하에서는 자본의 안정성과 수익성이 중요시되며, 특히 자본의 효율적인 운용과 조달형태는 향후 기업이 건전한 재무구조를 형성하는데 중요한 고려사항임을 알 수 있다.

오정훈(1998)의 연구에 의하면 일반적으로 도산한 기업은 대부분 매출, 자산 등 외형적인 성장에도 불구하고 수익성이 크게 악화되어 금융비용 부담이 늘어나는 등 전형적인 '거품성장'의 형태를 보이고 있었다는 것이다. 그러나 97년도에 도산이 양산된 것은 전형적인 경기침체에 기인한 탓도 없지 않으나 근본적으로는 과도한 차입, 특히 단기차입에 의존하여 수익성보다는 확장위주의 경영을 해왔던 것이 가장 근본적인 원인이라고 하였다. 그리고 이러한 기업의 구조적 부실이 경기침체에 따른 수익성 악화, 금융시장 불안으로 인한 금융권의 여신회수 등과 맞물리면서 도산의 급증을 야기한 것이다. 그리고 IMF후에 도산한 기업은 IMF전에 도산한 기업보다 훨씬 재무구조가 양호함에도 도산하는 사례가 많았는데 이 중 상당수가 단기적인 유동성 부족에 기인한 경우라고 하였다.

한국신용평가(1999)는 가장 영향이 큰 IMF전의 도산원인으로 자기자본의 부족을 제시하였으며 경기불황 및 업계부진, 수요의 감소, 다양한 거래처 확보의 태만, 과대투자, 무리한 기업확장, 시설의 노후, 보수 등을 두 번째 원인으로 지적하였다. IMF후의 도산원인으로 계열사수가 급증하여 수익성보다는 외형성장위주로 무리하게 사

업을 확장하였고 차입경영으로 인한 자본구조 취약, 자금조달 및 운용상의 문제점, 매출채권·재고자산 회전기간의 장기화에 따른 유동성 부족을 지적하였다.

김건우(1999)는 우리나라 기업의 도산원인을 규명하고, 응답자의 응답을 통하여 도산방지책을 도출하고자 도산한 104개 기업에 대해 설문조사를 하였다. IMF전의 도산원인은 기업구조 및 투자요인이 제 1요인이고 다음으로 경기·자금요인, 재무요인, 시장요인 등의 순으로 나타났으나, IMF후의 도산원인은 경기·자금요인이 제 1요인이고 다음으로 시장요인, 판매요인, 기업구조 및 투자요인 등의 순이라고 하였다. IMF후에 도산한 기업들은 기업내부 구조와 투자이사결정의 잘못보다는 IMF라는 외적환경변화로 인한 경기침체 및 자금사정 악화로 도산하였다고 주장하였다.

이상의 연구자들의 연구결과를 종합하면 다음과 같다. IMF전에 도산한 기업은 단기지급능력을 보유하고 있었음에도 투자에 따른 만성적인 수익성 악화와 금융비용 증가로 부채상환능력과 이기지급능력이 부족하여 도산하였다. 즉 주된 도산원인으로 잘못된 기업구조와 무리한 투자에 따른 수익성 악화를 들 수 있다. 반면, IMF후에 도산한 기업은 수익회복능력이 있었음에도 불구하고 외환위기로 인한 경기침체와 금융시장 붕괴에 따른 채권금융회사의 자금상환 압박으로 도산하였다. 즉 주된 도산원인으로 경기침체와 단기 지급능력 부족을 들 수 있다. 그리고 IMF후에 도산한 기업의 재무구조는 IMF전에 도산한 기업보다 양호하였다.

### 3. 기업도산예측모형에 관한 선행연구

#### 3.1. 기업도산예측 활용을 위한 인공지능경망

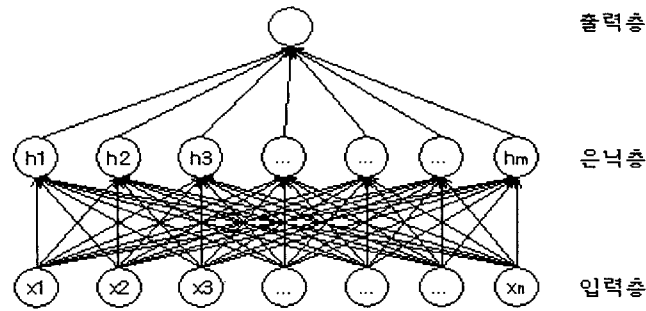
##### 3.1.1. 다층퍼셉트론의 인공지능경망 모형

Rosenblatt의 퍼셉트론은 학습가능한 층이 하나만 존재하기 때문에 단층퍼셉트론이라고도 불린다. 그러나 이것은 선형분리 가능한(Linear separable) 문제에만 적용할 수 있기 때문에 Exclusive-OR 문제를 해결할 수 없다는 한계가 있다. 이러한 한계를 벗어나는 방법은 인공지능경망의 입·출력층 사이에 은닉층을 부가하여 여러 계층의 구조를 갖도록 하는 것이다. 그런데 이러한 계층구조는 은닉층에 연결된 가중치 조정방법을 개발하는 데 많은 어려움이 따른다는 한계를 지녔으나, 1980년대 중반 PDP그룹에 의해 다층퍼셉트론을 백프로파게이션 학습알고리즘에 의해 학습시킬 수 있다는 것이 밝혀진 후, 이 알고리즘은 가장 널리 사용되는 모형이 되었다.

다층퍼셉트론은 입력층과 출력층 사이에 하나 이상의 은닉층(Hidden layer)이 존재하는 모형으로 <그림 1>에서와 같은 계층구조를 갖는다. 즉 다층퍼셉트론은 입력층→은닉층→출력층 순의 방향으로 연결되어 있는 전방향(Feedforward) 네트워크이다. 지도학습의 대표적인 예인 백프로파게이션 알고리즘은 전이함수의 미분값을 요구한다. 따라서 미분이 가능한 비선형함수인 시그모이드 함수를 노드내의 전이함수로 사용한다.

다층퍼셉트론에서 일반적으로 시그모이드 함수를 사용하는 이유는 신경망모형에 비선형성(Nonlinearity)을 도입하기 위함이다. 비선형성이 없다면 다층퍼셉트론의 능력은 은닉층이 존재하지 않는 단층퍼셉트론의 능력과 다를 바가 없게 된다. 그 이유는 입력을 선형으로 변형시키고 그 결과를 다시 선형으로 변형시키면 결과는 다시 선형이 되기 때문으로, 다층퍼셉트론의 사상(Mapping) 능력은 바로 비선형성에서 오는 것이라 할 수 있다.

<그림 1> 다층퍼셉트론의 아키텍처



다층퍼셉트론은 하나 이상의 은닉층을 두고 각 노드의 입출력 특성에 비선형함수를 활용함으로써 Minsky & Papert(1969)가 신경망의 한계점으로 비판한 선형분리 문제를 해결할 수 있게 되었다. 즉, 층의 개수가 증가할수록 퍼셉트론이 형성하는 의사결정 구역의 특성은 더욱 정교해지게 되는데 은닉층이 없는 경우에는 선형분리, 은닉층이 하나인 경우에는 오목한 개구역(또는 폐구역)을 형성하고, 은닉층이 두 개 이상인 경우에는 이론상 모든 형태의 구역을 형성할 수 있다(Lippman, 1987).

다층퍼셉트론에서 일반적으로 사용하는 백프로파게이션 알고리즘은 목표 출력값과 인공신경망의 출력값 사이에 오차의 제곱합을 최소화 하기 위해 최급하강법(Gradient-descent Method)을 사용하고 있는데, 이것은 전방향 계산과 역방향 계산으로 이루어진다. 전방향 계산은 주어진 입력값에 대해 인공신경망의 출력값을 계산하는 과정이며, 역방향 계산은 전방향 계산에 의해 얻어진 출력값과 목표출력값의 차이, 즉 오류를 출력층으로부터 입력층으로 역전파시켜가며 연결가중치들을 조정하는 과정이다. 백프로파게이션이라는 이름이 붙여진 이유도 여기에 있다(이재규 외, 1998).

다층퍼셉트론을 이용할 때 관심을 갖는 것은 일반화 능력이다. 다층퍼셉트론에서 일반화란 학습 중에 보지 못했던 데이터, 즉 학습용 표본과는 독립적인 검증용 표본에 대해 낮은 오차율을 보이는 능력을 말한다. 다층퍼셉트론의 오차를 최대한 낮추는 방향으로 학습을 지속할 경우 다층퍼셉트론의 일반화는 오히려 떨어지는데 이러한 현상을 과도적합(Overfitting)이라 한다. 이러한 현상이 발생하는 까닭은 인공신경망의 학습시간이 길어지면 표본에 내재된 독립변수와 종속변수의 관계를 일반화하는 것이 아니라 학습용 표본 자체를 기억하는 현상, 즉 잡음까지 학습하게 되기 때문이다(성시중, 1997).

### 3.2. 인공신경망을 이용한 도산예측모형의 선행연구

Odom & Sharda(1990)는 인공신경망과 다변량판별분석 기법으로 개발한 각각의 도산예측모형의 예측력을 비교하여 인공신경망 모형이 우수함을 증명하였다.

이건창(1993)은 도산예측에 대한 기존의 통계학적인 방법론이 갖는 한계를 극복하고 보다 실용적이고 효과적인 도산예측기법을 제안하기 위해서 인공지능기법을 사용하였으며 예측력 비교대상 모형은 다변량판별분석 모형, ACLS(ID3를 일반화시킨 귀납적 학습방법), 인공신경망 모형을 사용하였다. 인공신경망 모형은 다변량판별분석을 통해 선정된 재무비율을 이용하여 개발한 인공신경망 I 모형, 귀납적 학습방법을 통하여 선정된 재무비율을 이용하여 개발한 인공신경망 II 모형으로 세분하였다. 각 연구모형의 도산예측력 순위는 인공신경망 II 모형이 가장 우수하였고 다음으로 인공신경망 I 모형, 귀납적 학습방법 모형, 다변량판별분석 모형 순이다.

Boritz & Kennedy(1995)는 도산예측모형의 예측력이 변수선정, 표본추출방식, 인공신경망 유형 등의 변화에 따라 어떠한 영향을 미치는지를 파악하였다. 그들의 예측력 비교대상 모형을 상세히 분류하면 여섯 가지 네트워크 유형의 인공신경망 모형, 세가지 유형의 다변량판별분석 모형, 로짓분석 모형, 프라빗분석 모형, 기타(BP, OET) 등이다. 연구결과는 Ohlson(1980) 및 Altman(1968)의 연구에 사용된 변수를 결합하여 개발한 모형이 각 개별 모형보다 예측력이 우수하게 나타났다. 그러나 인공신경망 모형의 예측력은 다변량판별분석 모형, 로짓분석 모형, 프라빗분석 모형에 비해 월등히 차이가 있는 결과를 보이지 못하였고 각각의 변수 군에 따라 우수한 예측력을 보인 모형도 달랐다.

Suh & Kim(1996)은 인공신경망 모형을 이용하여 모형의 네트워크 크기, 아키텍처(Architecture), 표본구성, 입력변수 등을 기준으로 도산예측력을 비교하였다. 그들은 도산-부실기업간의 도산예측이 도산-정상기업간의 도산예측보다 어렵다는 것과 은닉층 노드수의 변화에 따라 도산 예측력에 차이가 있다는 것을 보여 주었다. 또한, 학습용 표본과 검증용 표본의 도산기업 비중은 제1종 및 제2종 오류에 영향을 미친다는 것을 알 수 있었으며 도산예측모형 개발시 우선적으로 고려되어야 할 사항으로 도산기업의 학습용 표본과 검증용 표본간에는 적절한 배분이 이루어져야 하며, 예측력을 높이기 위해서는 연구표본을 동질인 집단(Homogeneous data)으로 구성할 것을 주장하였다.

Back et al.(1996)은 다변량판별분석 또는 로짓분석을 통해 추출한 변수를 이용하여 개발한 인공신경망 모형의 도산예측력보다 GA(Genetic Algorithm)분석을 통해 추출한 변수를 이용하여 개발한 인공신경망 모형의 도산예측력이 더 우수한지를 검증하기 위해 다섯 개 유형의 모형을 개발하였다. 분석결과, 인공신경망 모형(GA를 통해 추출한 변수를 사용하여 개발한 신경망모형)이 가장 높은 예측력을 보였다.

Jo & Han(1996)은 도산예측력이 상대적으로 우수한 모형을 개발하기 위하여 개별모형보다는 CBFs, 인공신경망 모형, 다변량판별분석 모형을 결합한 모형을 개발하였다. 분석결과, 인공신경망 모형의 적중률은 82.48%로 다른 두 개의 방법보다 우수하였다. 또한 각각의 모형을 결합한 모형의 적중률은 90.8%로서 가장 우수한 결과를 보였으며 결합모형의 활용가능성을 보여 주었다.

Leshno & Spector(1996)는 자료의 양(data span), 학습기법, 학습회전수(number of iterations)에 따라 다양한 인공신경망 모형간의 도산 예측력을 비교하고, 또한 다변량판별분석 모형의 예측력과의 비교하였다. 예측력을 비교한 결과, 인공신경망 모형(도산 1, 2, 3년 전 표본을 모두 포함하여 우수한 결과를 보인 네트워크 구조로 100번의 회전을 통해 개발한 모형)의 총 평균 오분류율은 9.05%로서 상대적으로 우수한 예측력을 보였다.

성시중(1997)은 인공신경망 모형의 일종인 다층퍼셉트론(MLP: Multi-Layer Perceptrons)의 이론적 연구를 토대로 일반화 능력을 높일 수 있는 부실예측 모형을 개발하고 전통적으로 기업부실예측에 가장 많이 사용하는 다변량판별분석 모형을 구축하여 두 모형간의 예측력을 비교하였다. 분석결과를 보면 인공신경망 모형의 예측력이 판별분석보다 우수한 것으로 나타났다.

Tam & Kiang(1992)은 도산한 은행을 표본으로 개발한 선형판별분석 모형, 로짓분석 모형, kNN 모형, ID3 모형, 인공신경망 모형의 예측력을 비교하여 인공신경망 모형이 타 모형에 비해 우수한 것으로 확인하였다. 또한, 인공신경망 모형은 설명력 차원에서는 제한이 있음에도 불구하고 다른 모형에 비해 예측력이 높고 모형의 견고성(Robustness)도 높아서 시간에 따라 데이터가 추가되어도 새로운 사례에 적응력을 보인다는 점에서 기존 도산예측모형의 대안으로 활용될 수 있다고 주장하였다.

Fanning & Cogger(1994)는 로짓분석, 인공신경망(다층퍼셉트론) 그리고 연구자가 개발한 또 다른 인공신경망의 일종인 GANNA 모형을 개발하여 예측력을 비교하였다. GANNA와 로짓모형은 거의 비슷한 결과를 보였으며 다층퍼셉트론의 경우에는 두 모형에 미치지 못하는 예측력을 보였다.

<표 2> 인공신경망을 이용한 기업도산예측 선행연구

연구자	표본 대상기간	표본구성	분석방법	연구결과
Odom & Sharda(1990)	'75~'82	도산 65 정상 64	NN,MDA	NN이 상대적으로 우수한 예측력을 보임
Tam & Kiang(1992)	'85~'87	도산 59 정상 59	MDA,로짓분석 NN,kNN, ID3	NN이 상대적으로 우수한 예측력을 보임
이건창(1993)	'79~'92	도산 83 정상 83	MDA,ACLS NN I, NNII	NNII가 상대적으로 우수한 예측력을 보임
Boritz & Kennedy(1995)	'71~'84	도산 171 정상 6,153	NN(6가지), MDA(3가지), 로짓분석,프라빗 분석,BP, OET	특별히 우수한 예측력을 보인 기법은 없음
Suh & Kim(1996)	'90~'91 '89~'91	도산 68 정상 191 부실 191	NN (도산:정상기업) (도산:부실기업)	· 은닉층 노드수의 변화에 따라 예측력이 상 이함 · 도산-부실기업간 도산예측이 도산-정상기 업간 도산예측보다 어려움
Back et al.(1996)	'86~'89	도산 37 정상 37	MDA, 로짓분석 NN	GA를 통해 추출한 변수를 이용하여 개발한 NN의 예측력이 가장 우수
Jo & Han(1996)	'91~'93	도산 271 정상 271	CBFs, NN, MDA 결합모형	NN의 도산예측력은 82.48%이며 결합모형은 90.8%로 모형의 결합가능성 제시함
Leshno & Spector(1996)	'84~'88	도산 44 정상 44	NN, MDA	· 도산 1,2,3년 전 표본을 모두 포함한 경우 예측력은 상대적으로 우수함 · 상당히 향상된 네트워크 구조 또는 학습회 전수의 증가는 과도적합 초래 가능성 · NN의 오분류율이 MDA 보다 낮음
성시중(1997)	'93~'95	도산 80 정상 80	NN, MDA	NN의 예측력이 MDA보다 우수한 결과를 보임

\* NN : 인공신경망, MDA : 다변량판별분석, ID3·ACLS·CBFs : 귀납적학습방법,

### 3.3. 도산예측모형 개발기법간의 비교

본 연구의 분석기법인 다변량판별분석, 로짓분석 그리고 인공신경망 기법의 장·단점을 보면 다음과 같다.

다변량판별분석의 장점은 대중에게 널리 알려져 그 기법 자체의 신뢰성, 안정성이 높다는 것이다. 그리고 사용방법이 간편하고 결과에 대한 해석이 인공신경망 기법에 비해 쉬우며, 개발한 모형에 영향을 미치는 변수는 물론 그 영향의 정도를 파악할 수 있다. 또한 분석에 소요되는 시간이 인공신경망 기법에 비해 짧다. 그러나 기본가정으로서 독립변수들이 정규분포여야 한다는 점, 종속변수에 의해 범주화되는 집단들의 분산 및 공분산행렬이 동일해야 한다는 점이 방법론상 해결해야 할 선행 과제이다. 이견창 외(1994)에 의하면 가정의 위배에서 오는 문제점이 커서 효과적인 판별분석을 위해서는 기본가정에 대한 검증이 선행되어야 한다고 주장하고 있다. 또한 인공신경망 기법처럼 새로운 유형의 사례를 기존모형에 적용할 때 학습시킬 수 있는 메카니즘이 없다는 것이 단점이다.

로짓분석의 장점으로 판별분석처럼 독립변수들이 정규분포를 이루어야 하고, 종속변수에 의해 범주화되는 집

단들이 분산 및 공분산행렬이 동일해야 한다는 가정이 필요없다는 점이며 특이한 관찰치로 인해 모수추정에 미치는 영향이 적어 일반적인 선형회귀식의 추정에 비해 모형의 견고성이 높다. 또한, 로짓분석도 판별분석처럼 사용방법이 간편하고 결과 해석이 쉬워 공공성이 높으며 모형에 영향을 미치는 변수 파악도 가능하고 분석에 소요되는 시간도 인공신경망에 비해 짧다. 그러나 로짓분석의 결과 값은 확률로 표시되며 그 확률은 표본대상을 판별점을 기준으로 특정집단(도산, 정상)에 분류시키는 기준이 되므로 그 확률이 로지스틱 함수(S 자형)를 취한다는 가정이 필요하다. 그리고 판별분석과 마찬가지로 새로운 사례에 대한 학습 메카니즘이 없다는 것이 단점이라 하겠다.

위에서 기술한 바와 같이 두 기법 모두가 도산예측모형 개발에 있어 큰 차이는 없는 편이나 인공신경망 기법은 그 특성이 상이하다.

인공신경망의 장점은 모형의 견고성(Robustness)과 학습성(Learnability)을 들 수 있다(Lippmann, 1988). 견고성이란 인공신경망 자체의 특징인 오류허용성(Fault tolerance)인데 이는 특정 몇몇 노드에 오류가 발생하여도 인공신경망의 전체적인 기능에 큰 영향을 미치지 않는 것이다. 따라서 이러한 특징은 급격히 변화하는 환경이나 예측치 못했던 환경에서 인공신경망이 안정적으로 기능하는데 큰 역할을 한다는 것이다. 그리고 학습성이란 인공신경망이 주어진 학습 입력력 자료를 통해 숨겨진 규칙성(Regularities)을 찾아내는 과정을 의미한다. 이러한 학습을 통해 학습표본의 특성이 각 노드에 저장되는데 이러한 저장 자체가 지식베이스에 해당된다. 또 다른 장점으로서 인공신경망은 선형함수 뿐 만 아니라 비선형함수도 가정할 수 있고, 다변량판별분석 및 로짓분석과는 달리 새로운 유형의 사례에 대한 학습이 뛰어나 환경변화에 대한 적응력이 우수하다는 것이다. 그러나 인공신경망 모형은 모형의 일반성, 상황적응성 등이 높은 반면, 판별분석과 로짓분석과는 달리 추론과정에 대한 설명이 불분명하며 모형에 영향을 미치는 변수도 파악할 수 없어 분석결과에 대한 해석이 상대적으로 어렵다. 또한 입력변수를 시스템 자체에서 제공할 수 없으며 분석에 드는 시간이 많이 소요된다는 것이 단점이라 하겠다. 위의 내용을 요약하면 <표 3>와 같다.

<표 3> 도산예측모형 개발 기법별 장·단점

기법	장점	단점
인공신경망	· 모형의 견고성, 학습성, 상황적응성, 일반화가 상대적으로 높음	· 모형의 추론과정 설명 불투명 · 모형에 영향을 미치는 변수파악 불가 · 사용방법 및 결과해석의 어려움 · 신경망시스템내에서 입력변수 선정불가 · 분석에 많은 시간 소요
로짓분석	· 사용방법이 간편하고 해석용이 · 판별분석의 가정 불필요 · 판별분석보다 견고성이 높음 · 모형에 영향을 미치는 변수 파악가능	· 선택확률이 로지스틱 함수(S자형) 가정 · 모형의 학습메카니즘 부재
다변량 판별분석	· 사용방법이 간편하고 해석용이 · 모형에 영향을 미치는 변수 파악가능	· 독립변수가 정규분포이어야 함을 가정 · 분류집단의 공분산행렬이 동일함을 가정 · 모형의 학습메카니즘 부재

도산예측에 있어 인공신경망 기법은 로짓분석과 다변량판별분석에 비해 신뢰성과 안정성이 부족하기 때문에



널리 활용되지 못한 것이 현실이다. 그러나 기존연구(Back et al., 1996; Leshno & Spector, 1996; 이진창, 1993)에서 살펴보았듯이 인공신경망 모형은 예측력 및 상황적응력 측면에서 MDA 모형이나 로짓분석 모형보다 우수하였다.

## 4. 연구방법론

### 4.1. 표본수집

본 연구의 표본은 IMF후에 도산한 기업, 즉 '97년 10월 이후부터 '02년 1월말까지 한국증권거래소에 상장된 기업 중 제조업체를 모집단으로 하였고 다음과 같은 조건을 만족시키는 것을 대상으로 추출하였다.

- ① 도산의 개념적 정의 : 부도발생, 당좌거래정지, 회사정리절차개시, 화의절차개시 등
- ② 업종 : 제조업(건설업, 유통업 제외)
- ③ 도산 전 3개년간의 재무제표 및 기타 회계자료를 입수할 수 있는 기업
- ④ 금융회사가 아닌 기업

본 연구에서 연구표본을 상장회사로 한정시킨 이유는 상장회사가 기업규모나 경제활동 영역측면에서 국민경제에 미치는 영향이 지대하여 도산예측에 대한 사회적 실익이 크다는 점과 상장회사의 재무자료는 비상장회사 재무자료에 비해 수집이 용이하고 객관적이며 신뢰성이 높기 때문이다.

업종을 제조업으로 제한한 것은 상세업종마다 회계처리 방식이 다양하여 이로 인해 표본의 동질성이 떨어져 도산예측을 향상에 좋지 않은 영향을 미치므로 이를 예방하기 위함이다(Suh & Kim, 1996). 또한 유통업 및 건설업도 비록 제조업종에 속하나 타 제조업체에 비해 회계처리방식이 상이하여 도산예측을 향상에 부정적인 영향을 미칠 우려가 있어 본 연구의 표본에서 제외하였다.

도산기업에 대응되는 정상기업은 쌍대표본 추출방식으로 다음의 기준을 모두 만족시키는 기업으로 추출하였다.

- ① 도산기업의 도산 당시는 물론, 2002년 12월 현재에도 정상적으로 영업활동을 하고 있는 상장회사
- ② 98년 6월부터 2002년 12월 현재까지 워크아웃제도를 적용받지 않은 기업
- ③ 한국증권거래소 산업분류기준의 제조업종에 속하는 기업(건설업, 유통업 제외)
- ④ 도산기업의 도산 3년 전 자산과 매출액이 유사한 기업

이를 근거로 도산기업과 정상기업을 각각 82개, 총 164개를 연구표본으로 삼았다.

표본구성을 보면 학습용 표본은 97년부터 99년까지 도산·정상기업 각각 66개로서 총 132개이며, 검증용 표본은 2000년부터 2002년 1월말까지의 도산·정상기업 각각 16개로서 총 32개이다. 이처럼 검증용 표본을 학습용 표본기간 이후로 제한한 이유는 학습용 표본으로 개발한 모형에 미래 자료인 검증용 표본으로 예측력을 검증하여 그 모형의 일반화를 높이기 위함이다. 이는 학습용 표본과 동일한 기간의 검증용 표본에 대하여 예측의 정확성을 검토하는 기존의 분석방법과는 달리하는 것이기도 하다. 연구표본을 간략히 도표화하면 <표 4>과 같다.

<표 4> 연구표본의 구성

구분	97~99			00~02.1			합계
	도산	정상	소계	도산	정상	소계	
학습용 표본	66	66	132	-	-	-	132
검증용 표본	-	-	-	16	16	32	32
합계	66	66	132	16	16	32	164

본 연구에서는 도산을 예측하기 위한 재무정보로 도산 전 3년간 재무제표를 사용하였는데 그 이유는 기업도산이라는 것이 1년 또는 단시일내에 발생하기 보다는 어느 정도의 기간을 두고 발생할 가능성이 높기 때문이다. 연구표본의 재무자료는 한국신용평가정보에서 제공하는 기업정보(KIS-LINE)를 통해 확보하였다.

#### 4.2. 변수선정

본 연구에서 사용할 변수는 도산예측관련 기존문헌(이진창 외, 1994; Lee et al., 1996)에서 사용된 변수와 금융감독원(2002)에서 지정한 감사보고서상의 경영지표 중 기업도산을 잘 나타내 주고 자료의 입수가 용이하며 객관적이고 통계처리가 가능한 재무비율로서 다음과 같은 선정기준에 의해 추출하였다.

① 기존 국내·외 연구문헌들을 검토한 결과 네 번이상 사용된 변수로서 도산기업과 정상기업의 차이를 잘 나타내주는 비율

② 금융감독원에서 지정한 2002년부터 감사보고서상에 기재해야 하는 경영지표

③ 자료의 입수가 용이하고 객관적이며 통계처리가 가능한 재무비율

위 기준에 의해 선정된 재무비율은 당좌비율, 총부채현금흐름비율, 총자산, 총자산대비당좌자산, 총자산매출액비율, 총자산대비용전자본, 총자산대비용동자산, 총자산대비용이익잉여금, EBITDA이자보상배율, 매출액순이익율, 유동비율, 총자산현금흐름비율, 총자산순이익율, 차입금의존도, 부채비율, 이자보상배율, 매출액영업이익율, 자산회전율 등 18개이다. 모형에 영향을 미치는 의미있는 변수를 선별하기 위해 t-검정 후, 5% 유의수준에서 유의한 결과를 보인 재무비율을 연구에 사용할 최종적인 변수로 지정하였다.

또한 본 연구에서는 MDA 모형, 로짓분석모형 그리고 인공신경망 모형을 개발하며 특히, MDA 모형을 개발하기 위해서 다음과 같은 기본가정을 설정하였다.

첫째, 연구대상기간 뿐만 아니라 미래에 대해서도 도산·정상기업 표본의 모집단이 안정성을 가지고 있다는 점이다. 이는 일정기간을 대상으로 예측모형을 설계하고 이 모형을 이용하여 미래기간에 대한 예측을 실시하기 위한 기본가정이다.

둘째, 도산·정상기업 모집단의 판별점수의 분포는 정규분포를 이룬다는 가정으로 모집단에서 추출된 학습용 표본의 판별점수도 각각 정규분포를 이룬다는 가정이다.

셋째, 변수들은 각각 정규분포를 따른다.

넷째, 제1종 오류비용과 제2종 오류비용은 동일하다.

#### 4.3. 연구방법

좋은 네트워크(Well-generalized network)를 갖춘 인공신경망 모형을 개발하기 위해서는 인공신경망 유형, 네트워크의 크기, 네트워크의 형태, 학습알고리즘 그리고 전이함수 등이 모두 고려되어야 한다. 본 연구에서는

연구의 안정성을 높이기 위해 인공신경망 모형의 유형, 네트워크 토폴로지, 학습알고리즘 그리고 전이함수를 기존 문헌(이건창 외, 1994; Lawrence, 1993)에서 자주 사용되어 왔으며 현재 유행하고 있는 것으로 사용하려 한다. 네트워크의 크기는 명확한 기준이 없어서 여러 가지 방법들이 제시되고 있는데 그 방법들을 요약하면 아래와 같다.

은닉층의 노드수를 표본대상과 동수로 하여 네트워크에 적용시킬 경우 네트워크 학습에 상당히 부정적인 영향을 미칠 것이며, 반대로 노드수를 적게 할 경우에는 학습용 표본의 학습이 충분히 이루어지지 않는다. 이에 Lawrence(1993)는 최적의 은닉층 노드수 지정방법 중 대표적인 두 가지를 제시하였다. 그 중 하나는 노드수를 다양하게 제시하면서 각각의 네트워크를 학습시켜 최고의 값을 갖는 네트워크를 선정하는 방법이다. 다른 방법은 처음에는 노드수를 적게 하여 네트워크를 학습시켜 본 후, 학습이 정상적으로 이루어지지 않을 경우 다시 노드수를 하나씩 부가하며 네트워크를 학습시켜 가장 우수한 예측력을 보이는 네트워크를 선정하는 것이다.

전자는 시간이 많이 소요된다는 점과 어느 시점의 네트워크가 최적의 노드수를 갖는 때인지를 알기가 어렵다는 문제점을 지니고 있다. 이러한 문제에 직면하지 않기 위해서는 처음부터 노드수를 적게 하여 네트워크에 적용되 필요할 때마다 노드수를 하나씩 부가하여 최적의 노드수를 찾는 방법인 후자가 적합하다고 볼 수 있다.

이러한 측면에서는 시작 시점의 노드수를 어느 정도로 할 것인가 라는 문제가 발생하는데 일반적으로 입력노드수와 출력노드수를 합산한 수의 반(1/2)이상 정도이거나 학습용 표본수의 5~10%로 한다. 이와는 달리 처음부터 많은 수의 노드부터 시작하여 조금씩 노드수를 줄여가면서 최적의 네트워크를 선정하는 방법도 있다.

표본의 수는 네트워크의 크기와 비례하므로 네트워크가 크면 클수록, 즉 은닉층의 노드수가 증가할수록 학습용 표본의 수도 증가하므로, 통상 학습용 표본의 수는 적어도 은닉층의 노드수에 학습용 표본의 톨러런스(Training tolerance)의 역수를 곱한 값보다 많게 사용하고 있다.

좋은 네트워크를 위한 주요 요인의 하나로서 네트워크의 크기를 고려해야 하는데 이는 은닉층의 수와 직접 관련이 있다. 통상 네트워크의 구조는 하나의 은닉층을 갖는 네트워크로 세 개의 계층, 즉 입력층, 은닉층, 출력층으로 구성된다. 일반적으로 인공신경망 연구의 학자들은 작은 네트워크 구조를 권고하고 있다. 만일 두 개 이상의 은닉층을 갖는 네트워크를 설계한다면 학습하는데 많은 시간이 소요될 뿐만 아니라 은닉층 수의 증가에 따른 많은 표본이 필요하다는 문제점이 발생된다. 따라서 연구방법이 은닉층 수와 특별히 관계있는 것이 아니라면 한 개의 은닉층을 갖는 네트워크를 개발하는 것이 연구자체 뿐만 아니라 연구자에게도 시간적으로나 경제적으로도 절약된다(Lawrence, 1993).

이상의 내용을 바탕으로 본 연구에서는 다음의 방법으로 인공신경망 모형을 개발하고자 한다.

인공신경망 모형의 유형은 다층퍼셉트론이며 네트워크 크기는 은닉층이 하나인 계층구조를 갖는다. 네트워크의 형태(Topology)는 전방향 네트워크이며 학습알고리즘은 백프로파게이션을 사용한다. 은닉층의 노드수는 6개부터 39개까지 하나씩 부가하여 총 34개의 네트워크를 개발한다. 이 중 최고의 예측력을 나타내는 네트워크를 추출하여 그 네트워크의 예측력과 MDA 모형 및 로짓분석모형의 예측력을 비교하고자 한다.

## 5. 실증분석결과

### 5.1. t-검정 결과

기업도산을 예측하기 위해서는 우선 도산기업과 정상기업을 구별해 주는 변수를 선정해야 한다. t-검정은 도산기업과 정상기업 두 표본간의 분석대상인 재무비율 평균이 같다는 귀무가설의 유의성을 통계적으로 검정하는데 사용된다(남주하, 1998). 이에 기존문헌(이건창 외, 1994; Lee et al., 1996)에서 사용된 18개 재무변수 중 도산기업과 정상기업의 차이를 잘 나타내는 변수를 추출하기 위해 t-검정을 실시하였다.

도산 전 3년간 모두 통계적으로 유의한 차이를 나타내는 재무비율은 부채비율, 차입금의존도, 총자산매출액비율, 총자산순이익율, 총자산이익잉여금비율, 매출액순이익율, 이자보상배율, EBITDA이자보상배율, 총자산현금흐름비율, 총부채현금흐름비율, 총자산회전율 등 11개이다. 그러나 유동비율, 매출액영업이익율은 도산 1년 전에만 유의한 결과를 보였다. 이에 이들 두 개의 재무비율을 포함하여 총 13개를 입력변수로 사용하고자 한다.

본 연구에서 사용되는 입력변수를 유형별로 분류하여 보면 안정성 비율로는 유동비율, 부채비율, 차입금의존도 등 3개가 있으며, 수익성 비율로는 총자산매출액비율, 총자산순이익율, 총자산이익잉여금비율, 매출액순이익율, 매출액영업이익율, 이자보상배율, EBITDA이자보상배율 등 7개가 있고, 현금흐름 비율로는 총자산현금흐름비율, 총부채현금흐름비율 등 2개가 있으며, 활동성 비율로 총자산회전율이 있다.

<표 5> t-검정 결과

번호	변수명	통계적 유의성
1	당좌비율	.904
2	총자산	.660
3	총자산당좌자산비율	.318
4	총자산유동자산비율	.185
5	유동비율	.026**
6	부채비율	.000***
7	차입금의존도	.000***
8	총자산매출액비율	.000***
9	총자산순이익율	.000***
10	총자산이익잉여금비율	.000***
11	매출액영업이익율	.017**
12	매출액순이익율	.000***
13	이자보상배율	.001***
14	EBITDA이자보상배율	.002***
15	총자산현금흐름비율	.000***
16	총부채현금흐름비율	.000***
17	총자산운전자본비율	.304
18	총자산회전율	.000***

주) \*\*\* : 유의수준 0.01, \*\* : 유의수준 0.05

## 5.2. 다변량판별분석을 이용한 도산예측

다변량판별분석 가정 중에 중요한 것은 모든 독립변수가 정규분포를 따른다는 것이며 이견창(1993)에 의하면 독립변수를 정규분포로 전환하는 방법으로 Trimming, Truncation, Log-normality를 들고 있다. 본 연구에서는 Berry & Treigueiros(1991)의 연구에 따라 독립변수인 재무비율을 로그값으로 변환하였다. 단, 음(-)의 값을 갖

는 재무비율은 통계적으로 처리할 수 없어 로그값 변환에서 제외하였다.

본 연구에서는 총 18개 재무비율 중 t-검정 결과, 유의한 변수를 모형개발에 이용하여 윌크스 람다(Wilks' Lambda)를 기준으로 주요 변수를 선정하는 단계적 선택법을 실시하였다.

<표 6>은 다변량판별분석을 이용하여 학습용 표본을 대상으로 도산을 예측한 결과이다. 도산예측에 영향을 미치는 주요 변수는 세 개이며 판별점 및 판별식은 다음과 같다.

판별점 :  $-44.613 - (-54.517) = 9.904$

판별식(점수) :  $(52.057 - 47.374) * \text{총자산매출액비율} + (0.116 - 0.188) * \text{총자산현금흐름비율} + (0.06027 - 0.03867) * \text{부채비율}$   
 $= 4.683 * \text{총자산매출액비율} - 0.072 * \text{총자산현금흐름비율} + 0.0216 * \text{부채비율}$

기업도산을 판별하기 위해서는 위에서 추정된 판별식에 표본기업의 재무자료를 적용하여 판별점수를 계산하며 판별점수가 판별점보다 크면 정상기업으로 분류되고, 판별점보다 작으면 도산기업으로 분류된다.

<표 6> MDA 모형에 영향을 미치는 주요 변수

설명변수	시기	추정계수(Fisher's)	
		정상기업	도산기업
총자산매출액비율	도산2년전	52.057	47.374
총자산현금흐름비율	도산1년전	-0.116	-0.188
부채비율	도산1년전	0.06027	0.03867
상수		-54.517	-44.613

학습용 표본을 이용하여 개발한 MDA 모형의 판별식에 검증용 표본을 적용하여 모형의 예측력을 산출한 결과는 <표 7>과 같이 평균 78.125%의 예측력을 보이고 있으며 <표 8>와 같이 제1종 오류는 12.5%이고 제2종 오류는 9.375%이다.

<표 7> MDA 모형을 이용한 예측력 분석결과

(단위: 개, %)

구분	예측		정상기업	도산기업	계	예측력
	실제	예측				
학습용 표본	정상기업	50	16	132	78.05	
	도산기업	13	53			
검증용 표본	정상기업	13	3	32	78.125	
	도산기업	4	12			

<표 8> 검증용 표본의 오분류율 및 예측력

(단위: %)

판별점	제1종 오류율	제2종 오류율	예측력
9.904	12.5	9.375	78.125

5.3. 로짓분석을 이용한 도산예측

로짓분석은 다변량판별분석과는 달리 독립변수들이 정규분포이어야 한다는 가정이 필요 없다. 본 연구에서는 t-검정 결과 유의한 변수를 모형개발의 입력변수로 사용하였으며 각 변수의 극단치 제거를 위해 표준오차의 2배에서 벗어난 표본을 분석대상에서 제외시켰다. 그리고 모형에 영향을 미치는 주요 변수를 파악하기 위해 단계적 선택법(Stepwise Method)을 사용하였다.

<표 9>는 로짓분석을 이용하여 학습용 표본을 대상으로 도산을 예측한 결과로서 모형에 영향을 미치는 주요 변수 여섯 개가 산출되었으며 판별식은 아래와 같다.

<표 9> 로짓분석모형에 영향을 미치는 주요 변수

설명변수	시기	모수추정치(B)	통계적 유의성
총자산매출액비율	도산3년전	0.030	.019**
총자산이익잉여금비율	"	△0.082	.023**
유동비율	"	0.014	.011**
총자산매출액비율	도산2년전	△0.045	.003***
총부채현금흐름비율	도산1년전	△0.074	.006**
부채비율	도산1년전	△0.042	.002***
상수		1.699	

주) \*\*\* : 유의수준 0.01, \*\* : 유의수준 0.05

$$\text{판별식} : 1.699 + 0.030 \cdot \text{총자산매출액비율(도산3전)} - 0.082 \cdot \text{총자산이익잉여금비율(3전)} + 0.014 \cdot \text{유동비율(3전)} - 0.045 \cdot \text{총자산매출액비율(2전)} - 0.074 \cdot \text{총부채현금흐름비율(1전)} - 0.042 \cdot \text{부채비율(1전)}$$

가장 우수한 예측력을 보인 모형을 파악하기 위해 상기의 판별식을 검증용 표본에 적용하여 <표 10>과 같이 판별점을 0.1부터 0.8까지 적용시키면서 제1종 오류와 제2종 오류를 비교하였다. 비교결과 판별점이 0.29인 지점이 제1·2종 오류가 가장 낮으며 평균 예측력도 81.30%로서 가장 우수한 값을 보이고 있다.

<표 10> 판별점 이동에 따른 검증용 표본의 오분류율 및 예측력

(단위: %)

판별점	제1종오류율	제2종오류율	예측력
0.1	9.375	18.75	71.875
0.2	9.375	12.5	78.125
<b>0.29</b>	<b>9.375</b>	<b>9.375</b>	<b>81.30</b>
0.3	12.5	9.375	78.125
0.4	12.5	9.375	78.125
0.5	21.85	-	78.125
0.6	25.0	-	75.0
0.7	28.1	-	71.9
0.8	37.5	-	62.5

<표 11>은 <표 10>에서 가장 우수한 예측력을 보인 모형에 대해 분석한 결과이다. 학습용 표본으로 개발한 모형의 예측력은 82.55%를 보였다. 그 모형에 미래 자료인 검증용 표본을 적용한 결과, 예측력은 81.30%를 나타내고 있다.

<표 11> 로짓분석모형을 이용한 예측력 분석결과

(단위: 개, %)

구분	예측		정상기업	도산기업	계	예측력
	실제					
학습용 표본	정상기업	53	13	132	82.55	
	도산기업	10	56			
검증용 표본	정상기업	13	3	32	81.30	
	도산기업	3	13			

#### 5.4. 인공신경망을 이용한 도산예측

본 연구에서 개발한 인공신경망 모형의 설계는 다음과 같다. 인공신경망의 유형은 다층퍼셉트론을 사용하였으며 네트워크의 크기는 세 개층이다. 네트워크의 형태(Topology)는 전방향(Feedforward)네트워크이다. 학습알고리즘은 백프로파게이션(Backpropagation)을 사용하였으며 은닉층의 노드수는 학습용 표본의 5% 이상으로서 6개~39개이다. 즉, 최종 선정된 네트워크는 노드수 6개부터 39개까지 1개씩 부가하여 만든 34개의 네트워크 중 가장 우수한 예측력을 보이는 것으로 하였다.

가장 우수한 예측력을 보인 네트워크를 선정하기 위해 은닉층의 노드수를 6개부터 시작하여 1개씩 증가시켜 입력노드 수와 동수인 39개까지 총 34개 네트워크의 예측력을 비교하였다. <표 12>는 은닉층 노드수에 따른 예측력으로서 34개 네트워크 중 예측력이 상대적으로 높은 네트워크를 위주로 기술하였다.

<표 12> 은닉층의 노드수에 따른 오분류율 및 예측력

(단위: %)

노드수	제1종오류율	제2종오류율	예측력
6	9.38	6.25	84.38
9	6.25	9.38	84.37
<b>10</b>	<b>6.25</b>	<b>3.13</b>	<b>90.62</b>
11	12.5	3.13	84.32
<b>12</b>	<b>6.25</b>	<b>3.13</b>	<b>90.62</b>
13	6.25	6.25	87.50
17	9.38	3.13	87.51
20	12.50	3.13	84.38
29	9.38	6.25	84.38
30	9.38	3.13	87.51
33	9.38	3.13	87.51
35	12.50	3.13	84.38
39	9.38	3.13	87.51

<표 13> 인공신경망 모형을 이용한 예측력 결과

(단위: 개, %)

구분	예측		정상기업	도산기업	계	예측력
	실제					
검증용 표본	정상기업		15	1	32	90.62
	도산기업		2	14		

<표 12, 13>에서 알 수 있듯이 은닉층 노드수가 10개 및 12개인 네트워크의 예측력이 90.62%로 가장 우수한 결과를 보였다. 은닉층의 노드수가 많으면 네트워크 구조가 복잡해져 네트워크를 학습시킬 때 많은 시간이 소요되어 네트워크 구조는 은닉층 노드수가 적으면서 예측력이 높을수록 좋다고 할 수 있다. 이에 따라 예측력 비교에 사용될 가장 적절한 네트워크는 은닉층 노드수가 10개인 네트워크라 할 수 있다. 따라서 이 네트워크의 구조는 입력층 노드수 39개, 은닉층 노드수 10개, 출력층 노드수 1개로 구성(39-10-1) 되었으며 이 모형의 제1종 오류는 6.25%이고 제2종 오류는 3.13%으로 나타났다.

### 5.5. 각 모형의 예측력 비교

이상과 같이 학습용 표본을 모형개발의 학습자료로 사용하고 검증용 표본을 모형의 예측자료로 활용하여 MDA 모형, 로짓분석모형, 인공신경망 모형의 도산예측력을 살펴보았다.

본 연구의 목적은 모형간의 예측력 비교이다. 그러나 과연 예측력 비교를 퍼센트(%)로 표시하는 것이 통계적으로 타당한가라는 의문이 생길 수 있다. 본 연구에서 취한 방법론을 살펴보면 2000년부터 2002년 1월까지의 자료를 검증용 표본으로 하여 각 모형에 동일하게 적용한 후 예측력을 비교하였기 때문에 검증용 표본 수는 동일하다고 할 수 있다. 따라서 모형간의 예측력 비교측정 단위를 퍼센트로 하여도 그 결과는 통계학적 의미로 볼 때 전혀 문제가 되지 않는다. 이는 실제로 도산예측에 관한 기존연구(이건창 외, 1994; Lee, 1996)에서도 확인할 수 있다.

<표 14> 각 모형의 오분류율 및 예측력 비교

(단위: %)

구분	제1종오류율	제2종오류율	예측력
인공신경망 모형	6.25	3.13	<b>90.62</b>
로짓분석 모형	9.375	9.375	<b>81.30</b>
MDA 모형	12.5	9.375	<b>78.125</b>

<표 14>와 같이 예측력을 기준으로 각 결과를 비교하여 보면, 인공신경망 모형의 예측력이 로짓분석모형 및 MDA 모형의 예측력보다 우수한 것으로 나타났으며 로짓분석모형의 예측력이 MDA 모형의 예측력보다 우수한 결과를 보였다. 이상의 결과를 종합하면 인공신경망 모형 -> 로짓분석모형 -> MDA 모형의 순으로 예측력이 우수하다고 할 수 있다.

본 연구는 미래 환경변화에 대한 각 모형의 적응력을 검증하기 위하여 검증용 표본을 학습용 표본보다 미래 기간에서 추출하였다. <표 14>와 같이 학습용 표본으로 개발한 각각의 모형을 검증용 표본에 적용한 결과, 인



공신경망 모형의 예측력이 상대적으로 우수한 결과를 보여 인공지능경망 모형은 미래환경에 대한 적응력이 높다고 할 수 있다. 또한 인공지능경망 모형이 미래환경에 대한 적응력이 높다는 것은 모형의 견고성(robustness)이 높다는 것으로 인공지능경망에 기초한 도산예측모형은 급격히 변화하는 환경이나 예측하지 못했던 환경에서도 안정적으로 기업의 도산여부를 예측할 수 있다는 장점을 지니고 있다고 할 수 있다.

#### 5.6. IMF후 기업도산에 미치는 주요 요인

MDA 모형 및 로짓분석모형을 분석한 결과, IMF후 기업도산에 미치는 주요 재무비율로는 <표 15>와 같이 유동비율, 부채비율, 총자산매출액비율, 총자산이익잉여금비율, 총자산현금흐름비율, 총부채현금흐름비율이 밝혀졌다. 그러나 인공지능경망 모형은 도산예측모형에 어떤 변수가 어느 정도 영향을 미치는지를 파악할 수 없다는 단점을 지니고 있어 주요변수를 확인할 수 없다.

<표 15> IMF후 기업도산에 영향을 미치는 주요 변수

	로짓분석모형	MDA 모형
변수	유동비율	부채비율
	부채비율	총자산매출액비율
	총자산매출액비율	총자산현금흐름비율
	총자산이익잉여금비율	
	총부채현금흐름비율	

특히, 1997년 10월부터 1999년 12월말까지 도산한 기업들은 당시 경기침체로 인해 소비가 감소하고 판매실적이 부진해짐에 따라 수익이 감소되고 유동성이 부족하게 되어 금융회사로부터 높은 금리에 자금을 조달받게 되었다. 이는 차입금 증대에 따른 회사의 부채가 증가되는 결과를 가져왔고 또한 당시 경기침체에 따른 채권금융회사의 자금상환압박과 결부되어 도산하게 되었음을 보여주고 있다.

## 6. 결론

본 연구는 기업도산예측에 대한 기존의 통계학적인 방법론이 갖는 한계를 극복하고 보다 실용적이고 효과적인 도산예측 기법을 제안한다는 차원에서 인공지능경망 기법을 이용하였다. 또한, IMF라는 경제환경 변화로 인해 연구표본을 IMF후에 도산한 기업으로 하고 예측력 향상을 위해 동질의 자료로서 상장회사 중 제조업종으로 한정하였으며 모형의 성과를 검증하기 위하여 MDA 모형, 로짓분석모형, 인공지능경망 모형의 예측력을 비교하였다.

학습용 표본으로 개발한 각 모형에 검증용 표본을 적용하여 산출한 도산예측력을 분석한 결과, 로짓분석모형의 예측력 81.30%, MDA 모형의 예측력 78.125%에 비해 인공지능경망 모형의 예측력이 90.62%로 가장 우수함을 알 수 있다. 이는 기존연구 결과(Back et al., 1996; Leshno & Spector, 1996; 이진창, 1993)와 같이 인공지능경망 모형의 예측력이 가장 우수하며, 도산예측연구의 방법론으로서 인공지능경망 모형의 유용성이 높음을 보여주고

있다. 그러나 인공신경망 모형은 통계적 분석방법의 장점인 변수추출기능이 없으므로 독자적으로 사용하는 것보다 통계적 분석방법과 병행하는 것이 보다 효율적이라 할 수 있다.

그리고 실증분석 결과, IMF후 기업도산에 미치는 주요 재무비율로는 유동비율, 부채비율, 총자산매출액비율, 총자산이익잉여금비율, 총자산현금흐름비율, 총부채현금흐름비율임이 밝혀졌으며 이들은 안정성, 수익성, 현금흐름 비율과 관련된 것이다. 이러한 결과를 통해 볼 때, IMF후에 도산한 기업은 재무구조가 불안정하고 기업내부의 유동성 부족 등으로 단기지급능력이 악화되어 도산하였음을 미루어 짐작할 수 있다.

위와 같은 연구결과가 시사하는 바는 첫째, 검증용 표본을 학습용 표본기간 이후로 하여 분석한 결과를 보면 인공신경망 모형의 예측력이 가장 우수하다는 점이며 인공신경망 기법을 활용한 도산예측 모형은 미래 환경변화에 적응력이 높다고 할 수 있으므로 실무활용 등의 일반화 가능성을 내포하고 있다.

둘째, 금융회사 입장에서는 인공신경망 모형을 기업신용위험 상시평가시스템 및 조기경보시스템의 근간이 되는 기업신용등급 평가모형으로 활용할 수 있을 것이다. 또한 기업 입장에서는 자사의 신용도 및 거래처 신용도 파악 등을 통해 경영활동 의사결정에 활용할 수 있을 것이다.

본 연구의 한계점과 향후 연구방향을 요약하면 다음과 같다.

첫째, 선행 연구에서 개별 모형보다 결합모형의 예측력이 우수하다는 결과가 나왔고 도산예측연구의 추세가 결합모형인 점을 감안하여 볼 때, 본 연구는 인공신경망 모형, 로짓분석모형, MDA 모형 등 각각의 개별 모형의 예측력 산정 및 비교에 한정되어 있다는 한계를 지닌다.

둘째, 연구표본은 IMF후에 도산한 업체를 대상으로 하였으나 변수선정은 IMF와 관계없이 기존문헌 등에서 주로 쓰인 변수를 사용하였는데 IMF후에 도산한 업체에 영향을 미치는 주요 변수를 추출하여 도산예측모형을 개발하는 것이 필요하다.

따라서 앞으로의 연구방향으로는 첫째, 본 연구의 표본이 제조업체인 상장회사를 중심으로 하였으나 향후에는 타업종 또는 비상장회사를 중심으로 개발하는 것도 도산예측모형의 공공성을 위해 의미있는 연구가 될 것이다.

둘째, 도산예측모형을 이용한 특정기업의 예측력과 각 금융회사의 여신심사역이 그 기업에 대해 평가한 결과를 상호비교한 후, 그 차이를 분석하여 개발한 모형을 보완·모형의 일반화를 기하여야 할 것이다.

셋째, 연구표본을 IMF후에 도산한 기업으로 한정할 이유는 IMF전·후의 기업도산 원인이 다소 상이하기 때문이다. 이에 변수선정도 본 연구와 같이 기존문헌에서 주로 사용한 변수보다는 오히려 IMF후의 기업도산 원인과 관련된 변수를 이론적·통계적으로 추출하여 사용하는 것이 바람직하다고 볼 수 있다. 향후에는 IMF후에 도산한 기업의 도산원인을 입력변수로 사용이 필요하다.

## 참고문헌

- 금융감독원, 기업신용위험 상시평가 관련 보도자료, 2001.2~2002.10
- 김건우(1999), "한국기업의 도산원인," 경영학연구, 제28권, 제3호, pp.771-798.
- 남명수(1998), "IMF전·후의 기업부실원인 변동분석," 한국상장회사협의회 연구조사자료.
- 남주하(1998), "기업의 부도원인과 부도예측모형 분석: IMF 전후 기간을 중심으로," 금융연구, 제12권, 제2호, pp.77-107.

- 성시중(1997), "인공신경망을 이용한 기업부실예측", 고려대학교 박사학위논문.
- 오정훈(1998), "IMF이전과 이후 부도기업의 특징", LG주간경제.
- 이건창(1993), "기업도산예측을 위한 통계적 모형과 인공지능 모형간의 예측력 비교에 관한 연구 : MDA, 귀납적 학습방법, 인공신경망", 한국경영과학회지 제18권 제2호, pp.57-81.
- 이건창, 김명중, 김혁(1994), "기업도산예측을 위한 귀납적 학습지원 인공신경망 접근방법: MDA, 귀납적 학습방법, 인공신경망 모형과의 성과비교," 경영학연구, 제23권, 제3호, pp. 109-144.
- 이재규 외(1998), 「전문가시스템 : 원리와 개발」, 법영사.
- Altman, E. I.(1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, Vol.23, No.4, pp.589-609.
- Back, B., Laitinen, T., and Sere, K.(1996), "Neural Networks and Genetic Algorithms for Bankruptcy Predictions," *Expert Systems With Applications*, Vol. 11, No. 4, pp.407-413.
- Boritz, J. E. and Kennedy, D. B.(1995), "Effectiveness of Neural Network Types for prediction of Business Failure," *Expert Systems With Applications*, Vol. 9, No. 4, pp.503-512.
- Fanning, K. M. and Cogger, K. O.(1994), "A Comparative Analysis of Artificial Neural Networks Using Financial Distress Prediction," *International Journal of Intelligent Systems in Accounting, Finance, and Management*, Vol. 3, pp.241-252.
- Jo. H. and Han, I.(1999), "Integration of Case-Based Forecasting, Neural Network, and Discriminant Analysis for Bankruptcy Prediction," *Expert Systems With Applications*, Vol. 11, No. 4, pp.415-422.
- Jo, H., Han, I. and Lee, H.(1997), "Bankruptcy Prediction Using Case-Based Reasoning, Neural Networks, and Discriminant Analysis," *Expert Systems with Applications*, Vol. 13, No. 2, pp. 97-108.
- Lawrence, J.(1993), "Introduction to Neural Networks: Design, Theory, and Application," California Scientific Software Press, 5th edition, January.
- Lee, K. C., Han, I., and Kwon, Y.(1996), "Hybrid Neural Network Models for Bankruptcy Predictions," *Decision Support Systems*, Vol. 18, pp.63-72.
- Leshno, M. and Spector, Y.(1996), "Neural network prediction analysis: The bankruptcy case," *Neurocomputing*, Vol. 10, No.2, pp. 125-147.
- Lippmann, R. L.(1987), "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*.
- Luther, R. K.(1998), "An Artificial Neural Network Approach to Predicting the Outcome of Chapter 11 Bankruptcy," *Journal of Business and Economic Studies*, Vol. 4, No.1, pp. 57-73.
- Odom, M. and Sharda, R.(1990), "A Neural Network Model for Bankruptcy Prediction," *Proceeding of the IEEE International Conference on Neural Networks*, Vol.2, pp. 163-168.
- Ohlson, J. A.(1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal Accounting Research*, Vol. 18, No.1, pp.109-131.
- Suh, Y. H. and Kim, J. B.(1996), "Current Artificial Neural Network Models For Bankruptcy Prediction," *Journal of Accounting & Business Research*, Vol. 4, pp.81-101.
- Tam, K. Y. and Kiang, M. Y.(1992), "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions," *Management Science*, Vol. 38, No. 7. pp. 926-947.
- Zhang, G., Hu, M. Y., Patuwo, B. E. and Indro, D. C.(1999), "Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis," *European Journal of Operational Research*, Vol. 116, No. 1, pp.16-32.