

부도예측용 인공신경망모형의 최적 입력노드 설계 :
연결강도판별분석 접근
(Design of Optimal Input Nodes in Artificial Neural Network
Models for Bankruptcy Prediction: Link Weight
Discrimination Analysis Approach)

이용규 (wklee@biho.taegu.ac.kr) 손동우 (sdw3296@hotmail.com)

대구대학교 경영학과
전화번호 : 053-850-6272
팩스 : 053-850-6239
경북경산시 진량면 내리리 15
(우편번호 712-714)

요약

인공신경망에 의해 부도예측을 하기 위해서는 여러 개의 재무비율을 입력변수 즉, 입력노드로 이용하는데, 이 가운데 적절한 입력노드를 선정하는 일은 예측력을 결정하는데 있어서 매우 중요하다. 본 연구에서는 새로운 입력노드 선정 휴리스틱을 제안하기 위하여 적절한 훈련이 끝난 인공신경망 모델에서 각 입력노드와 연결되는 가중치들의 합에 대한 절대값인 연결강도가 작을 경우 해당 노드는 출력값에 대한 설명력이 약할 것이라라는 연결강도판별 명제를 제시한다. 즉, 연결강도가 연결강도임계치 보다 작은 입력노드는 제거 대상으로 분류할 수 있을 것이고, 이들 노드를 제외한 입력노드는 그렇지 않은 경우보다 더 나은 예측력을 보여 줄 수 있을 것이다. 연결강도판별 명제를 실증적으로 입증하기 위해 본 연구에서는 연결강도판별 선처리 과정에 대한 방법론을 제안하고 제안된 방법론에 의해 부도예측을 실시하여 아무런 선처리를 거치지 않은 모형과 비교하였고, 또 기존의 입력변수 선정방식 중에 하나인 의사결정 트리 방식에 의한 입력변수 선정 모형과도 비교하여 더 나은 결과를 얻었다.

1. 서론

인공신경망 모형은 다변량 판별분석과 같은

통계적 기법과는 달리 독립변수인 입력노드에 대한 별도의 통계적 가정 없이도 사용할 수 있기 때문에 부도예측(Bankruptcy Prediction)문제를 해결하는데 널리 사용되어 왔고, 그 예측력이 기존의 통계적 기법 보다 우수한 것으로 알려지고 있다.

다른 응용분야에서와 마찬가지로 인공신경망에 의한 부도예측에 있어서 적절한 입력노드의 선정은 모형의 예측력을 높이기 위해 매우 중요하다. 일반적으로 부도예측을 하기 위해서는 여러 개의 재무비율을 입력변수인 입력노드로 이용하게 되는데 이 가운데는 부도예측에 유의한 영향을 미치지 못하는 것이 있을 수 있기 때문에 부적절한 변수를 입력노드에서 제거시키는 작업은 다변량 판별분석과 같은 통계적 기법에서나 인공신경망 모형에서도 중요한 연구이슈로 대두되고 있다.

본 연구에서는 부도예측용 인공신경망 모형의 최적 입력노드 설계를 위한 새로운 휴리스틱으로 연결강도판별분석 접근법(Link Weight Discrimination Analysis Approach)을 제안한다.

연결강도판별분석 접근법은 인공신경망 모형의 입력노드를 선정함에 있어서, 인공신경망 모형의 기본 메카니즘을 이용하는데 입력노드(Input Nodes)와 은닉노드(Hidden Nodes)간의 가중치(Weighted Value)가 적을, 다시 말해 연결강도가 약한 입력노드를 제거하는 선처리(Pre-processing) 과정을 거쳐서 최적의 입력노

드를 선정하고자 하는 방법이다.

본 연구에서는 연결강도판별분석 접근법의 타당성을 입증하기 위해 연결강도판별분석에 의해 선처리된 인공신경망 모형과 의사결정트리에 의해 선처리된 인공신경망 모형, 그리고 입력변수 선정을 위한 선처리 과정을 거치지 않은 순수 인공신경망 모형의 성과를 각각 비교·분석한다.

본 논문의 구성은 다음과 같다.

2절에서는 부도예측을 위한 통계적 기법과 인공지능 기법을 살펴보고, 인공신경망에서의 변수 선정을 위한 기법과 각 기법의 한계점을 논의한다. 3절에서는 본 연결강도판별분석 접근법의 이론적 배경과 방법론 그리고 그 예를 보이고, 4절에서는 연결강도판별분석 접근법에 의해 선처리된 인공신경망 모형과 선처리 없는 인공신경망 모형 및 의사결정트리에 의해 선처리된 인공신경망 모형의 성능을 비교분석하도록 한다. 마지막으로 5절에서는 본 연구의 기여도 및 향후 과제에 대해 논의한다.

2. 문헌고찰

2.1 부도예측 기법

인공신경망 모형이 발표 되기 이전까지 부도예측은 다변량 판별분석(Altman, 1968)이나 로짓회귀분석(Logit Regression), 프로빗분석(Probit Analysis) 등의 회귀분석(Regression) (Ohlson, 1980 ; Zmijewski, 1984)과 같은 통계적 기법이 주로 이용되어 왔었으나, 1980년대 이후부터 인공신경망 모형에 의한 부도예측에 관한 연구가 활발히 이루어지고 있다.

현재 부도예측에서 가장 널리 활용되고 있는 인공신경망 모형은 다층퍼셉트론(Multi-Layered Perceptron) 구조를 가지고 있으면서 역전파 알고리즘(Back Propagation Algorithm)에 의한 지도학습(Supervised Learning)에 기반하고 있는데, 기존의 다변량 판별분석과 같은 통계적 기법보다 우수한 예측력을 보이고 있는 것으로 알려지고 있다(조흥규, 1999 ; Tam and Kiang 1991 ; Fletcher, 1993 ; Tsukuda, 1994 ; Udo, 1993 ; Wilson, 1994). 다층퍼셉트론 이외에도 자기조직형상화지도(Self-Organization Features Map) (Serrano-Cinca, 1996 ; Chen, 1995)와 인공지능의 귀납적 학습 방법 중에 하나인 의사결정트리(Sung, et al, 1999 ; Messier and Hansen, 1988) 등도 부도예측에 적용하여 좋은 예측율을 보여 주고 있다.

2.2 부도예측용 인공신경망에서의 입력변수 선정

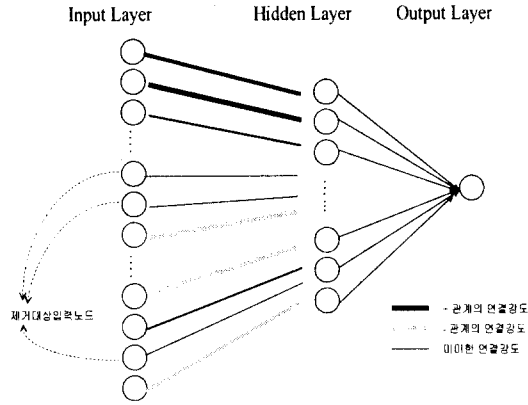
부도예측용 인공신경망의 입력변수 선정 기법으로 널리 알려진 것은 다변량 판별분석, 자기조직형상화 지도, 의사결정 트리 등을 들 수 있는데 이들 기법을 이용하여 선처리를 실시한 인공신경망 모형이 그렇지 않은 경우보다 높은 예측력을 보여 주고 있는 것으로 알려져 있다(Lee, et al, 1996 ; 조흥규, 1999).

그러나, 기존의 입력변수 선정기법은 나름대로의 한계점을 보유하고 있는데, 다변량 판별분석의 경우 독립변수들의 집단이 다변량 정규분포를 이루면서 공분산이 같아야 하는 가정이 필요하기 때문에 이에 대한 검증을 필요로 한다(채서일, 1999 ; Tam and Kiang, 1992). 의사결정트리의 경우 독립변수와 종속변수의 관계를 쉽게 설명할 수 있는 IF~THEN 규칙을 생성하는 장점이 있어서 결과를 쉽게 이해할 수 있지만(Messier and Hansen, 1988), 독립변수가 취할 수 있는 값이 많아질 경우에는 오류가 발생할 가능성이 많다는 약점을 가지고 있다(Berry and Linoff, 1997). 자기조직형상화 지도에 의한 변수 선정방법은 패턴인식이 없는 자율학습(Unsupervised Learning) 방법을 채택하고 있기 때문에 대상변수에 대한 설명력이 부족하다는 데에 한계가 있다.

3. 연결강도판별분석 접근법

3.1 연결강도판별 및 약체연결뉴런제거명제

[그림-1]에서 보는 바와 같이 학습이 끝난 상태의 인공신경망에서 각 입력노드와 연결되는 가중치들의 합은 해당 입력노드가 이 인공신경망의 출력값에 영향을 미치는 정도로 생각할 수 있다. 즉, 연결된 가중치의 합이 정(+)의 값을 가질 경우 해당 입력노드는 전체 출력값에 정의 영향을 미친다고 볼 수 있고, 부(-)의 값을 가질 경우는 부의 영향을 미친다고 볼 수 있으며, 그 값의 절대값이 미미할 경우에는 출력값에 대한 설명력이 약하다고 판단될 수 있다.



[그림-1] 연결강도판별분석 접근법의 개념

따라서 본 연구에서는 다음과 같은 연결강도 판별 명제와 약체연결뉴런제거 명제를 제안한다.

연결강도판별 명제: 적절한 훈련이 끝난 인공신경망 모델에서 각 입력노드와 연결되는 가중치들의 합에 대한 절대값이 작을 경우 해당 노드는 출력값에 대한 설명력이 약하다.

약체연결뉴런제거 명제: 연결강도판별치가 작은 입력노드를 제거한 인공신경망 모형은 그렇지 않은 경우보다 출력값에 대한 설명력이 높다.

3.2 연결강도판별 분석법

연결강도판별분석 접근법(Link Weight Discrimination Analysis Approach)은 인공신경망 모형에서 연결강도가 약한 입력노드는 전체 출력값에 미치는 영향이 미미하기 때문에(연결강도판별 명제) 입력노드와 연결되는 가중치의 합에 대한 절대값인 연결강도 판별식(Link Weight Discrimination)이 일정값인 연결강도 판별임계치(Link Weight Discrimination Cut off Value)보다 작은 경우 해당 입력노드를 제거(약체연결뉴런제거 명제)하여 최적의 인공신경망 구조를 찾으려는 입력노드 선정 휴리스틱이다.

연결강도판별분석 접근법에서는 일단 인공신경망을 학습시켜 최적의 인공신경망을 구한 다음 이 인공신경망에서 입력노드와 연결된 가중치의 합에 대한 절대값인 연결강도판별식이 연결강도판별 임계치보다 작은 입력노드를 제외하고 나머지 입력노드들만으로 새롭게 학습을 한다. 이와 같은 과정을 다음과 같은 경우가 나올 때까지 계속한다. 첫째, 연결강도판별식에

의한 값이 연결강도판별 임계치보다 작은 입력노드가 나오지 않거나 둘째, 인공신경망의 성능이 더 이상 개선되지 않거나 셋째, 반복회수가 일정 회수에 도달할 때까지 반복한다.

인공신경망의 학습방법은 연결가중치의 값을 무작위 방식에 의해 초기화하여 이 초기화한 연결가중치를 점진적으로 향상시키는 방식으로 전개된다. 따라서 인공신경망에 의한 학습결과는 항상 무작위성을 내포하게 되어 있다. 이러한 무작위성을 반영하기 위해 연결강도판별분석 접근법에서는 위에서 제시한 과정을 일정 회수만큼 반복하여 여기에서 얻어진 제거대상 입력노드의 빈도(Frequency)를 구하고, 제거대상으로 선정된 빈도수가 일정한 횟수 이상이 되는 입력노드들을 전체 입력노드에서 제거하고 남은 입력노드를 최종 입력노드로 한다.

이를 정형적인 방식으로 정리하면 다음과 같다.

n 개의 입력노드와 m 개의 은닉노드 및 1개의 출력노드를 갖는 인공신경망이 있다고 하자. 그리고 입력노드는 x_i 은닉노드는 h_j 출력노드는 z 라고 하자.

x_i : 입력노드 ($i = 1, 2, \dots, n$)

h_j : 은닉노드 ($j = 1, 2, \dots, m$)

z : 출력노드

또 학습이 완성되어 최적화 되었을 때의 입력노드 x_i 와 은닉노드 h_j 간의 가중치 값을 w_{ij} 라고 할 때 연결강도 판별식 y_j 를 다음과 같이 정의한다.

$y_j = |\sum w_{ij}|$ (모든 j 에 대해서)

제 0 단계) k 와 s 를 각각 1로 한다.

제 1 단계) 입력노드 x_i , 은닉노드 h_j , 출력노드는 z 로 이루어진 인공신경망에 의해 학습을 시행하여 최적화시킨다.

제 2 단계) 각각의 y_j 를 구하여 제거대상 변수목록에 첨부한다. 즉, 다음을 만족하는 y_j 의 D_{sk} 를 구한다

$D_{sk} = \{y_j : y_j < \alpha\}$ (α : 연결강도 판별임계치)

제 3 단계) $D_{sk} = \emptyset$ 이거나 오류율이나 오분류율이 필요이상으로 커지고 있다고 판단될 경우 제 4단계로 가고 그렇지 않을 경우 입력노드들 가운데서 D_{sk} 에 속하는 해당되는 x_j 를 제거하고 s 값을 1 증가시키면서 제 1 단계로 돌아간다.

제 4 단계) k 값이 실험예정회수 M 보다 작을 경우 k 의 값을 1 증가시키고 제 1 단계로 돌아가고 그렇지 않을 경우 중단한다. 위 알고리즘에서 나온 $\{D_{sk}\}$ 의 원소들의 빈도수를 조사하여 빈도수가 N 보다 큰 제거대상변수를 골라 이들을 제외한 나머지 변수들만을 본 처리의

입력노드로 선정한다.

3.3 연결강도판별분석 접근법의 예

다음은 본 연구에서 실험한 연결강도판별분석 접근법의 한 예를 보여주고 있다. $\alpha=0.1$, $M=10$, 즉 연결강도판별 임계치는 0.1, 라운드수는 10회로 하여 실험을 수행하였다. [표-1]은 1라운드에서 선정된 제거대상변수(D_1)와 그 가중치를 보여주고 있다. 1라운드의 첫 번째 수행 결과는 X10, X13을 비롯한 총 10개(D_{11}), 두 번째에는 X9, X15를 비롯한 14개(D_{21}), 세 번째는 X41, X50을 비롯한 5개(D_{31})의 제거대상변수가 선정되었다. 네 번째는 오분류율(Missclassification Rate)이 필요 이상으로 커졌기 때문에(0.24→0.38) 실험을 중단하였다.

[표-1] 라운드별 제거대상변수

반복 회수	제거 변수	가중치	반복 회수	제거 변수	가중치
D ₁₁	X10	0.086	D ₂₁	X23	0.013
	X13	0.074		X27	0.087
	X14	0.056		X35	0.031
	X25	0.013		X38	0.078
	X29	0.065		X43	0.005
	X32	0.09		X47	0.053
	X34	0.097		X52	0.032
	X36	0.068		X56	0.085
	X39	0.086		X70	0.064
	X59	0.003		X41	0.029
D ₂₁	X9	0.088	D ₃₁	X50	0.094
	X15	0.007		X55	0.023
	X18	0.017		X66	0.022
	X19	0.042		X67	0.048
	X21	0.039		총 29개 변수	

[표-2]는 10라운드에 걸쳐 반복회수를 최대 10회까지 실험한 결과, 제거대상변수로 선정된 변수들을 빈도수별로 나타내고 있다.

[표-2] 빈도수별 제거대상변수

빈도수	제거대상변수
10	X15, X35
9	X9, X14, X21, X23, X25, X29, X36, X56, X59, X67
8	X18, X27, X47, X62, X69, X73
7	X10, X19, X34, X39, X41, X52, X61, X66, X70
6	X31, X32, X33, X37, X38, X43, X46, X49, X50, X53, X55, X60
5	X13, X28, X42
2	X44

4. 비교분석

4.1 데이터 수집

본 연구에 이용되는 데이터는 한국신용평가(주)에서 제공한 총 150개 기업의 자료이며, 69개의 부도기업과 81개의 건전기업의 데이터로 구성되어 있다. 인공신경망 모형에서 학습 및 검증용으로 100개 기업의 데이터를 선정하여 이용하였으며, 100개 기업 가운데 60%를 학습(Training)에 사용하였고 40%를 학습에 대한 검증(Validation)에 사용하였다. 나머지 50개 기업의 데이터는 학습을 거친 인공신경망 모형의 부도예측 적중율을 확인(Scoring)하는데 사용하였다. 본 연구에서는 SAS에서 개발한 엔터프라이즈 마이너(Enterprise Miner)3.0을 이용하였다.

4.2 입력변수

인공신경망 모형의 입력노드는 43개의 재무변수를 이용했는데, 상세한 변수내역은 부록에 첨부하도록 한다.

4.3 실험방법

연결강도판별분석 접근법에 의해 선정된 입력변수를 인공신경망 모형에 적용하여 얻은 모형(NN+WL)과 인공신경망을 독립적으로 수행한 모형(NN), 의사결정 트리로 선처리한 후 인공신경망 기법을 적용한 모형(NN+DT)을 비교·분석한다.

NN+WL(x)는 3절에서 보여준 연결강도판별분석 접근법에 의해 선정된 입력변수들을 인공신경망 모형에 적용하여 실험을 실시하였다. 즉, 은닉층의 수를 1개로 고정시킨 후 학습을 통해서 모형을 최적의 상태로 만들고, 얻은 결과에서 연결강도판별 임계치인 α 를 0.1로 하고, 라운드수 M을 10으로하여, 나머지 변수를

다시 인공신경망 모형에 적용한다. 이와 같은 과정을 제거대상변수가 발생하지 않거나, 혹은 오분류율(Misclassification Rate)이 불필요하게 커지지 않을 때, 그리고 반복회수가 10을 넘어가지 않을 때까지 반복해서 라운드별로 제거대상변수를 가려내었다. 제거대상변수를 제거한 후 남은 변수를 입력변수로 하고 은닉층의 노드수는 입력층 노드의 수와 같게 하였다.

다음으로 NN+DT는 의사결정 트리를 이용하여 입력변수 선정을 위한 선처리를 거쳐서 선정된 입력변수들을 인공신경망 모형에 적용하여 실험을 실시했으며, NN+WL(x)의 경우와 같이 은닉층의 노드수를 입력층의 노드수와 같게 하였다.

NN은 43개의 변수를 그대로 이용하여 인공신경망 모형을 이용했으며 모형의 최적화를 위해서 은닉층의 수를 각각 3, 5, 10, 15, 20, 25, 30, 35, 40, 43, 45, 50으로 조정해 가면서 실험하고 가장 양호한 결과가 나오는 은닉층의 수를 선택하여 10회에 걸쳐 시드를 조정해 가면서 실험을 실시하였다.

4.4 실험결과

4.4.1 연결강도판별분석 접근법에 의해 선처리된 인공신경망 모형(NN+WL(x))

빈도수별로 제거대상변수를 정리하여 10회, 9회, 8회, 7회, 6회에 걸쳐서 제거대상으로 나타난 변수를 순서대로 누적하여 제거했으며, 그 결과 10회에 걸쳐서 제거대상변수로 선정된 변수는 X15, X35 두 개의 변수이고, 9회 제거대상 변수로 선정된 것은 X9, X14 등을 비롯한 10개의 변수, 8회 제거대상 변수로 선정된 변수는 X18, X47 등 6개 변수, 7회 제거대상 변수로 선정된 변수는 X10, X19 등 9개의 변수, 6회 제거대상 변수로 선정된 변수는 X31, X32 등 12개의 변수가 각각 선정되었다. 이 실험의 결과인 빈도수별 제거대상변수의 목록은 [표-2]에서 보여주고 있다. 위와 같이 빈도수별로 제거대상변수로 선정된 변수들을 누적하여 제거하면서 각각을 인공신경망 모형에 적용하였으며, 10회 제거대상변수로 선정된 변수들을 제거한 인공신경망 모형을 NN+WL(10), 10회와 9회 선정된 변수를 제거한 모형을 NN+WL(9), 10회, 9회, 8회 선정된 제거대상변수를 누적하여 제거한 인공신경망 모형을 NN+WL(8), 10회, 9회, 8회, 7회 선정된 변수를 누적하여 제거한 인공신경망 모형을 NN+WL(7), 10회부터 6회까지 선정된 제거대상 변수를 누적하여 전부 제거한 인공신경망 모형을 NN+WL(6)으로 각각 나타냈다.

각 모형에 대한 실험을 수행한 결과 NN+WL(6)모형의 예측적중율이 78%로 가장 높게 나타났다. [표-3]은 각 모형별, 반복회수별 예측적중율을 보여주고 있다.

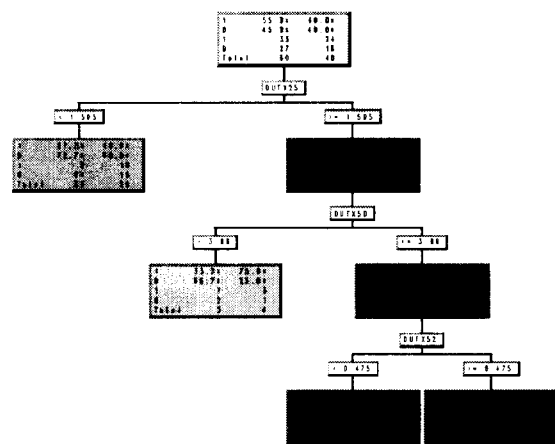
[표-3] 각 라운드와 모형별 예측적중률(%)

R	NN+WL (10)	NN+WL (9)	NN+WL (8)	NN+WL (7)	NN+WL (6)
1	56	66	68	60	68
2	64	60	64	68	66
3	58	62	62	56	72
4	56	64	56	58	70
5	60	66	62	60	66
6	62	58	60	58	68
7	62	60	64	68	60
8	58	68	64	60	70
9	60	58	66	70	74
10	58	66	60	60	78*

* : 최대 예측율

4.4.2 의사결정트리에 의해 선처리된 인공신경망 모형(NN+DT)

제거대상변수를 선정하기 위해서 이번에는 의사결정트리 기법을 이용하여 입력노드 선정을 위한 선처리과정을 진행하였다. 의사결정트리 기법에 의해 선정된 제거대상변수는 X9, X10, X13 등으로, 전체 43개의 변수 가운데 X25, X50, X52를 제외한 나머지 40개의 변수가 제거대상변수로 선정되었다. 의사결정트리에 의한 제거대상변수 선정 실험의 결과로 나타난 트리는 [그림-2]와 같다.



[그림-2] 의사결정트리에 의한 변수 선정

위의 실험결과에 근거하여 제거대상변수로 선정된 40개의 변수들을 제거하고, 입력변수로 선정된 입력노드는 X25, X50, X52였다. 이와 같이 선정된 입력노드들을 인공신경망 모형에

적용, 학습 후 스코어링한 결과는 32개 기업의 부도와 건전을 예측하여 예측적중율 66%로 나타났다.

4.4.3 순수 인공신경망 모형

인공신경망 모형을 독립적으로 이용했을 때의 결과는 다음과 같다. 50개의 스코어링 데이터로 확인하였으며, 은닉층 수별 예측적중율을 살펴보면 은닉층 수가 20개일 때 52%로 가장 낮은 예측적중율을, 은닉층의 수가 43개 일때가 각각 62%로 가장 높은 예측적중율을 보여주었다. 예측적중율이 높은 은닉층의 수 43개를 대상으로 시드를 조정하여 예측적중율을 다시 측정하였다. 10회에 걸쳐 시드를 조정하여 실험을 실시하였으며 그 결과, 62%가 가장 좋은 예측적중율로 나타났다.

4.5 실험결과분석

지금까지 실험을 통한 각 모형의 예측결과를 살펴보면, [표-4]에서 보는 바와 같이 연결강도판별분석 접근법에 의해 입력노드 선정과정을 거친 인공신경망 모형(NN+WL(x))에 의한 예측율은 78%, 의사결정트리에 의한 입력노드 선정과정을 거친 인공신경망 모형(NN+DT)의 예측율은 66%, 인공신경망 모형을 독립적으로 이용했을 때(NN)의 예측율은 62%로 각각 나타났다. 예측율의 수치로 그 결과를 비교해 보면 연결강도판별분석 접근법에 의한 방법이 인공신경망을 독립적으로 이용한 모형보다 우수한 예측력을 보여 주었고 의사결정트리에 의해 선처리 과정을 거친 인공신경망 모형 보다도 높은 예측율을 보여 주고 있다. 한편 이들 예측율에 대한 쌍체비교 Z-test를 수행한 결과는 연결강도판별분석 접근법에 의해 선처리된 인공신경망 모형의 예측율이 선처리를 거치지 않은 인공신경망모형에 비해 유의수준 5%에서 차이를 보여 주고 있고, 의사결정트리에 의해 선처리된 모형도 차이를 보여 주고 있음을 알 수 있다.

[표-4] 각 모형별 예측율

사용모형	NN+WL(x)	NN+DT	NN
예측율	78%	66%	62%
(쌍체비교 Z 값)	(-)	(1.4924**)	(1.6962*)

* : 유의수준 5%

** : 유의수준 10%

제거대상변수로 선정된 변수들 가운데, 안정성 관련비율인 자기자본비율(X44) 변수는 연결

강도판별분석 접근법에 의한 제거대상변수 선정시에도 제거빈도 2회를 나타내서, 본 연구에서 최적 모형으로 제시한 NN+WL(6)모형에서 최적입력노드로 선정되었다. 그 밖에도 제거빈도 5회를 나타낸 성장성 지표의 총자본 투자효율(X13), 수익성 지표의 손익분기점율(X28), 안정성 지표의 유동부채대 총자본(X42)과 같은 변수도 연결강도판별분석 접근법에 의한 부도 예측용 인공신경망 모형에서 예측적중율과 관련된 중요한 변수임을 알 수 있다.

5. 결론

본 연구는 인공신경망 모형에서 입력노드 선정을 위한 선처리 과정의 새로운 방법론으로서 연결강도판별분석 접근법을 제시했고, 연결강도판별분석 접근법에 의해 입력노드 선정과정을 거친 인공신경망 모형의 예측률이, 입력노드 선정을 위한 선처리 과정을 거치지 않은 순수인공신경망 모형이나 의사결정트리에 의한 입력노드 선정과정을 거친 인공신경망 모형보다 더 우수함을 입증하였다.

향후 연구과제로는 위해서는 첫째, 연결강도판별분석 방법론의 보완이 필요하다. 가령 본 연구에서 제시한 정지규칙을 좀 더 타당한 방식으로 개선하면서 정형적인 방식으로 정리할 필요가 있다. 둘째, 성능 개선을 위해서 NN+WL(x)의 각종 매개변수(Parameters)들에 대한 민감도 분석이 을 통해 다양한 방식으로 모형을 최적화 시킬 필요가 있다. 연결강도판별분석에서 사용하는 중요한 매개변수는 연결강도판별임계치, 제거대상 빈도수, 라운드 수 등을 들 수 있는데, 이들 매개변수의 값이 인공신경망의 성능에 미치는 영향을 규명하여 이들에 대한 최적화를 찾는 일이 연구 이슈로 제기될 수 있다. 셋째, 데이터 크기의 확장과 구성의 균형화(Balanced Data) 그리고 다른 방법론과의 비교 등을 통한 실험 결과의 일반성 확보다. 본 연구에서 제시한 휴리스틱이 일반화되기 위해서는 데이터셋을 확장하는 것 이외에 다양한 데이터 셋에의 적용을 필요로 한다. 또 샘플상에 부도와 건전 기업간의 비율을 균형화함으로써 학습상에 반영되는 정보의 양을 최적화할 필요가 있다. 또 다변량 판별분석과 같은 통계적 기법과 자기조직 형상화지도와 같은 다른 인공지능 기법들과의 성능비교를 필요로 한다. 넷째, 연결강도판별분석을 하나의 의사결정지원 시스템으로 구축하는 일이다. 선처리 과정과 본 처리과정을 일련의 시스템으로 구성하여 전 과정의 자동화는 물론이고 사용자와 적절한 인터

랙션을 할 수 있는 메카니즘을 제공함으로써 일종의 종합적인 데이터마이닝 도구로 발전시키는 것도 연구 이슈가 될 수 있을 것이다. 본 연구에서 제안한 연결강도판별분석 접근법을 부도예측 이외에도 채권평가(Bond Rating)와 같은 다른 분류(Classification)문제에 적용하는 것도 향후의 연구과제가 될 수 있을 것이다.

참고문헌

- Altman, Edward I., "Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy", *Journal of Finance*, Sep.1968, pp.589-609.
- Berry, Michael J. A. and Linoff, Gordon., *Data Mining Techniques: For Marketing, Sales, and Customer Support*, John Wiley and Sons, 1997
- Chen, S. P., Mangiameli and D. West, "The Comparative Ability of Self-Organizing Neural Networks to Define Cluster Structure", *Omega*, 23. 3, 1995, pp.271-279
- Fletcher, D. and Goss, E., "Forecasting with Neural Networks: an Application using Bankruptcy Data", *Information and Management*, 24. 3, 1993, pp.159-167
- Lee, Kun Chang, Han, Ingoo and Kwon, Youngsig, "Hybrid neural network models for Bankruptcy Predictions", *Decision Support Systems*, 18, 1996, pp.63-72.
- Messier, William F. Jr. and Hansen, James V., "Inducing Rules For Expert System Development: An Example Using Default and Bankruptcy Data", *Management Science*, Vol.34, No.12, December. 1988, pp.1403-1415.
- Ohlson, J. "Financial Ratios and The Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, 18. 1, 1980, pp.109-131.
- Serrano-Cinca, Carlos, "Self Organizing Neural Networks for Financial Diagnosis", *Decision Support Systems*, 17, 1996, pp.227-238.
- Sung, Tae Kyoung., Chang, Namsik and Lee, Gunhee, "Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy Prediction," *Journal of Management Information Systems*, Vol.16, No.1, Summer.1999, pp.63-85.
- Tam, Kar Yan And Kiang, Melody Y., "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions", *Management Science*, Vol.38, No.7, July.1992, pp.926-947.
- Tsukuda, J., and Baba, S. L., "Predicting Japanese Corporate Bankruptcy in terms of Financial Data Using Neural Network", *Computers and Industrial Engineering*, 27(1-4), 1994, pp.445-448
- Udo, G., "Neural Network Performance on the Bankruptcy Classification Problem", *Computers and Industrial Engineering*, 25(1-4).1993, pp.377-380.
- Wilson, Rick L. and Sharda, Ramesh, "Bankruptcy Prediction using Neural Networks", *Decision Support Systems*, 11. 1994., pp. 545-557.
- Zmijewski, Mafke, "Methodological Issues Related to the Estimation of Financial Distress Prediction Models", *Journal of Accounting Research*, Vol.22 1984,
- 조흥규, "도산예측을 위한 인공지능 방법과 통계적 방법의 통합 방법론," 한국과학기술원 (박사학위논문) 1999.
- 채서일, '사회과학 조사방법론', 학현사, 1999. , 2판.

부록 (사용변수 내역)

번호	지표	변수명	변수내역
1	규모(1)	x9	매출액
2	규모(1)	x10	자기자본
3	생산성(2)	x13	총자본투자효율
4	생산성(2)	x14	부기가치율
5	성장성(3)	x15	매출액증가율
6	성장성(3)	x18	재고자산증가율
7	성장성(3)	x19	총자산증가율
8	수익성(4)	x21	금융비용대 매출액비율
9	수익성(4)	x23	대출효율성계수
10	수익성(4)	x25	매출액순이익률
11	수익성(4)	x27	매출원가율
12	수익성(4)	x28	손익분기점율
13	수익성(4)	x29	순금융비용대 매출액비율
14	수익성(4)	x31	이자보상배율
15	수익성(4)	x32	자기자본순이익률
16	수익성(4)	x33	총자본경상이익률
17	수익성(4)	x34	총자본순이익률
18	안정성(5)	x35	고정장기적합율의 역
19	안정성(5)	x36	단기부채대 총차입금
20	안정성(5)	x37	당좌비율
21	안정성(5)	x38	매출채권대 매입채무
22	안정성(5)	x39	순운전자본비율
23	안정성(5)	x41	유동비율
24	안정성(5)	x42	유동부채대 총자본
25	안정성(5)	x43	유보액대총자산비율
26	안정성(5)	x44	자기자본비율
27	안정성(5)	x46	차입금의존도
28	안정성(5)	x47	총차입금대 매출액
29	추세(6)	x49	금융비용부담금율 증가율
30	활동성(7)	x50	매입채무회전율
31	활동성(7)	x52	순운전자본대 매출액
32	활동성(7)	x53	운전자금대 회전율
33	활동성(7)	x55	재고자산회전율
34	활동성(7)	x56	총자본회전율
35	현금흐름지표(8)	x59	(영업활동후 CF/차입금합계)*100
36	현금흐름지표(8)	x60	(영업활동후 CF/부채총계)*100
37	현금흐름지표(8)	x61	영업활동후 CF/금융비용
38	현금흐름지표(8)	x62	(현금영업이익/금융수입)/금융비용
39	현금흐름지표(8)	x66	(이자지급후 CF/부채총계)*100
40	현금흐름지표(8)	x67	이자지급후 CF/금융비용
41	현금흐름지표(8)	x69	(외부자금조달전 CF/차입금합계)*100
42	현금흐름지표(8)	x70	외부자금조달전 CF/금융비용
43	현금흐름지표(8)	x73	외부자금조달후 CF/금융비용