

Data Mining 기법들과 전문가들로부터 추출된 지식에 관한 실증적 비교 연구*

김광용*.손광기**.홍은선***

A Comparative Analysis for the Knowledge of Data Mining Techniques with Expertises

Gwang-Yong Gim*.Gwang-Ki Son**.On-Sun Hong***

요 약

본 연구는 여러 가지 Data Mining 기법들로부터 도출된 지식과 AHP를 이용하여 도출된 전문가의 지식을 사용된 정보의 특성에 따라 조사하고, 이러한 각각의 지식들을 중심으로 부도예측 모형을 설계한 후, 각 모형의 특성 및 부도예측력에 대한 실증적 비교연구에 그 목적을 두고 있다. 사용된 Data Mining 기법들은 통계적 다중판별분석모형, ID3 모형, 인공신경망 모형이며, 전문가 지식의 추출은 AHP를 사용하여 45명의 전문가로부터 부도와 관련하여 인터뷰 및 설문조사를 실시하였다. 특히 부도예측에 사용된 변수의 특성을 정량적 재무정보와 정성적 비재무정보로 나누어서 각 모형의 특성을 비교연구하였다. 연구결과 부도예측시 정성적정보의 중요성을 확인하였으며, 전문가의 지식을 기반으로 한 AHP 모형이 위험예측모형으로 사용될 수 있음을 실증적으로 보여주었다.

주제어 : AHP, Data Mining, 인공신경망, ID3, 부도예측, 정성적 정보

*본 논문은 1997학년도 숭실대학교 연구비 지원에 의해 연구되었음.

**본 논문은 1997학년도 광운대학교 연구비 지원에 의해 연구되었음.

* 숭실대학교, 경영학부 조교수

** 보험감독원, 경영학 박사

*** 광운대학교, 전자재료공학과 부교수

I. 서 론

본 연구는 여러 가지 Data Mining 기법들로부터 도출된 지식과 AHP(Analytic Hierarchy Process)를 이용하여 도출된 전문가의 지식에 대한 특성 및 각 지식을 기반으로 한 부도모형의 예측력에 대한 실증적 비교연구에 그 목적을 두고 있다. 사용된 Data Mining 기법들은 통계적 다중판별분석(DA) 모형, ID3 모형, 인공신경망(NN) 모형이며, 전문가 지식의 추출은 AHP를 사용하여 45명의 전문가로부터 부도와 관련하여 인터뷰 및 설문조사를 실시하였다. 특히 부도예측에 사용된 변수의 특성을 정량적 재무정보와 정성적 비재무정보로 나누어서 각 모형의 특성을 비교연구하였다.¹⁾

기존의 부도예측과 관련된 많은 연구에서는 주로 정량적(Numeric) 재무정보를 이용하여 통계적인 예측모형을 도출하거나 인공지능 기법을 활용한 모형의 도출을 많이 시도하였다. 그러나 실제로 현업에 종사하고 있는 전문가들의 설문결과, 대다수의 전문가는 계량적 모형으로부터 도출된 결과를 이용하기보다 전문가의 판단을 이용한 부도예측(Judgemental Forecasting)을 주로 하는 것으로 조사되었다(Bunn and Wright, 1991; Cerullo & Avila, 1975; Klein and Linneman, 1984). 부도예측을 측면에서도 전문가들의 판단에 기인한 부도예측율이 통계적모형보다 높은 경우도 많으며, 통계적모형과 전문가 지식을 혼합한 방법의 부도예측율이 통계적 모형이나 전문가의 판단보다 더 높게 나타난 연구결과도 많이 있었다

(Lawrence, 1985; Edmundson, 1988; Wolfe and Flores, 1990).

이러한 관점에서 볼 때 기존의 통계학적 접근뿐만 아니라 인공지능을 이용한 여러 가지 Data Mining 기법들로부터 도출된 지식과 전문가로부터 도출된 지식의 특성을 비교연구하고, 그러한 여러 가지 특성이 부도 예측력과는 어떠한 관계가 있는지 조사하는 것은 향후 Data Mining 기법의 바람직한 활용방안에 일조할 것으로 기대된다.

따라서 본 연구에서는 기업부도 예측변수에 대한 가중치를 계층화 의사결정법(AHP: Analytic Hierarchy Process, 이하 AHP라 함) 설문지를 이용한 인터뷰 방식을 통하여 실제로 부도예측 전문가가 생각하고 있는 각 변수의 부도예측에 관한 상대적 중요도(가중치)를 도출하고 Data Mining 기법들로부터 도출된 각 변수의 가중치와 비교함으로써 각 지식들의 특성을 비교하고자 한다. 또한 이러한 가중치를 이용한 AHP 부도예측모형을 설계한 뒤, 실제기업의 부도 또는 건전 자료와 비교하여 본 모형의 예측율을 검증하고, Data Mining 모형의 예측율과도 비교하여 각 방법으로부터 도출된 지식과 부도예측력과의 비교연구를 하고자 한다. 마지막으로 기존의 정량적인 재무정보에 주로 의존한 부도 예측 모형에서 탈피하여 정량적인 재무정보와 정성적인 재무정보를 같이 고려함으로써, 사용되는 정보의 형태에 따라 어떻게 지식이 활용되는 가를 살펴보고, 특히 부도예측에 있어 정성적 비재무정보의 중요성을 검증하고자 한다.

1) 이재식, 한재홍(1995)은 중소기업도산예측에 있어 비재무정보의 유용성을 검증한 바 있다

II. 국내 · 외 연구동향

2.1 여러 가지 Data Mining 기법들

〈표 1〉에서 보듯이, 1980년 중반부터 1980년 말까지 부도위험을 포함한 여러 가지 위험을 분류하고 예측하는데 LOGIT 또는 DA(Discriminant Analysis)같은 통계적 선형모형(Statistical Linear

Model)과 ID3같은 Rule Induction의 부도예측을 비교하는 연구가 비교적 많이 이루어져 왔다. 그후 〈표 2〉에서 보듯이 1990년 초반부터 인공신경망(Neural Network)을 나머지 두 방법과 비교하는 연구들이 이루어져 왔으며, 특히 많은 연구들이 인공신경망의 유효성을 보여주었다(김광용, 1996)

〈표 1〉 통계적선형모형 접근과 Rule Induction 의 실증적 비교

	통계적 선형모형	Rule Induction	적용문제
Minger (87)	Regression*	ID3*	Football Game (2 Cases)
Carter & Catlett (87)	Regression	ID3 & C*	Credit Assessment
Braun & Chandler (87)	DA	ID3*, Human	Stock Market
Messier & Hansen (88)	DA	ID3*, Human	Loan Default, Bankruptcy
Shaw & Gentry (90)	LOGIT	AQ15*	Business Loan
Liang et al. (90)	DA	ID3, CRIS*	Bankruptcy
Chung & Silver (92)	LOGIT	D3, GA*	Admission, Bidder
Cronan & Glorfeld (91)		ID3, RPA*	Mortgage lending
Arizne & Narasimha (94)	Regression*	ID3	Academic Performance

* 가장 높은 예측력을 보임

〈표 2〉 세가지 Data Mining 방법들의 실증적 비교

	Rule Induction	통계적선형모형	인공신경망	적용문제
Tam (91)	ID3	LOGIT, DA	NN*	Bank Bankruptcy
Shavlik et al. (91)	ID3		NN*	Simulation
Salchenberger et al. (92)		LOGIT	NN*	Thrift & Failures
Liang (92)	CRIS*, ACLS, PLS1	DA	NN*	Bankruptcy, Inventory
Liang et al. (92)	ID3	Probit	NN*	Inventory
Hansen et al. (92)	ID3*	Probit, LOGIT	NN*	Audit decision
Tam & Kiang (92)	ID3	LOGIT	NN*	Bank bankruptcy
Kattan et al. (93)	Human*, ID3, RPA*	DA	NN	Bank loan
Whalen & Gim (93)		LOGIT*	NN	Medical Malpractice
Gim (96)	ID3	DA	NN*	Auto Insurance
Gim & Whalen (96)	D3	LOGIT*	NN*	Bankruptcy

* 가장 높은 예측력을 보임

이러한 Data Mining 기법들과 관련된 연구들의 초점은 어떤 방법이 위험분류에 있어서 좋은 예측력을 갖는지에 대한 비교검증을 주로 하였다. <표 1, 2>에서 보듯이 현재까지도 어떤 Data Mining 기법이 예측력 측면에서 가장 우수하다고는 결정된 것은 없고 여러 가지 적용되는 문제의 특성에 따라 각기 다른 예측력을 보여왔다. 즉 각각의 방법들은 장단점을 갖고 있으며 각기 문제영역이 달라짐에 따라 다른 예측력을 보여주고 있다. 근래에는 각각의 방법들의 장점을 통합한 통합모형에 대한 많은 연구가 이루어지고 있다 (Kandal and Langholz, 1992; Whalen and Gim, 1996). 그러나 Hart(1992)가 지적했듯이 이러한 Data Mining 기법들의 결과가 위험예측력면에서는 유효할 수 있으나 실제 이러한 기법들의 결과를 이용하는 전문가의 검증없이 사용하는 것은 많은 문제를 야기할 수 있다. 따라서 실제 사용되는 측면에서 Data Mining 기법의 지식과 전문가의 지식과의 여러 가지 특성에 대한 비교연구가 선행되어야하나 이러한 연구는 많이 이루어지지 않았다.

2.2 전문가의 판단과 통계적 선형모형

인간의 의사결정에 있어 통계적 선형모형의 유용성(Dawes and Corrigan, 1974)과 그와는 반대로 인간 판단에 있어 통계적 선형모형의 부적절성(Zeleny, 1976)이 논의된 이후, 전문가의 판단과 통계적 선형모형의 위험예측력비교에 대한 연구는 여러분야에서 매우 활발히 진행되었다. 특히 기존의 많은 연구는 주어진 문제에 대한 위험예측에 있어 전문가 또는 초심자(Novice)의 개별적 판단(Separate Judgement)과 통계적 선형모형의 실증적 비교 (일명 'clinical vs statistical')가 주종을 이루었으나, 최근에는 통계적 선형모형의 결과를 이용한 인간의 혼합적 판단(Adjusted or Combined Judgement)과 통계적 선형모형과의 비교가 많이 이루어져왔다 (Bunn and Wright, 1991). 일반적으로 많은 연구에서 통계적 선형모형이 인간판단보다 위험예측력면에서 우수하거나 동등하였는데, 인간의 판단이 우수한 경우만을 따로 뽑아 <표 3>에 정리하였다 (김광용, 1997).

<표 3> 인간의 판단과 통계적선형모형의 실증적 비교

	인간판단	실험대상	적용문제	특징
Libby (75)	Seperate	Expert	Bankruptcy	여러 재무비율을 사용할 경우 인간판단 우수
Whitred & Zimmer (85)	Seperate	Expert	Bankruptcy	모델에서 사용하지 않은 정보를 사용한 경우 인간판단우수
Shepanski (83)	Seperate	Expert	Bank Loan	비선형적 특징이 있는 경우 인간판단우수
Daws et al. (89)	Seperate	Expert	laboratory	"Broken Leg" Que가있는 경우 인간판단 우수
Chalos (85)	Seperate	Expert	Credit Scoring	추가 정성적정보의 사용시 인간판단우수
Casey & Selling (86)	Seperate	Expert	Financial Distress	추가 정성적정보의 사용시 인간판단우수

	인간판단	실험대상	적용문제	특징
Reinmuth & Guerts (72)	Adjusted	Expert	Sales Forecast	통계적선형모형의 결과 이용 시 인간판단 우수
Mathews & Diamantopoulos (89)	Adjusted	Expert	Sales Forecast	통계적선형모형의 결과 이용 시 인간판단 우수
Huss (86)	Adjusted	Expert	Sales Forecast	통계적선형모형의 결과 이용 시 인간판단 우수
Wolfe and Flores (90)	Adjusted	Student	Forecasting Earning	통계적선형모형의 결과 이용 시 인간판단 우수

〈표 3〉에서 보듯이 인간의 판단이 우수한 경우는 통계적 선형모형의 도움없이 비선형적, 정성적 추가정보를 이용한 경우이며, 또한 통계적 선형모형의 도움을 받은 경우는 선형적, 정량적 자료를 이용한 경우가 인간의 판단이 우수하였다. 현재까지 이루어진 많은 연구는 인간의 총체적 판단(Holistic Judgement)과 여러 가지 통계적 선형모형과의 비교가 주종을 이루었다. 그러나 중요한 것은 인간판단이 통계적 선형모형보다 우수하나 열등하나 또는 통계적 선형모형을 이용한 인간판단이 우수하나 열등하나가 아니라, 각 방법의 특징 및 장, 단점을 비교, 연구하여 구조적인 상호관계(Structural Interaction)를 규명함이 우선되어야 한다 (Bunn & Wright, 1991). 특히 근래에는 통계적 선형모형 이외에도 인공신경망(NN)이나 ID3와 같은 인공지능(Artificial Intelligence)을 이용한 많은 Data Mining 기법들이 사용되고 있어 이들의 지식과 인간의 판단지식 사이에 어떤 관계가 있는지에 대한 비교연구도 많이 이루어져야 한다.

2.3 전문가의 지식을 도출하기 위한 AHP모형

본 연구에서 사용하고자 하는 계층화 의사결정법(analytic hierarchy process: AHP)은 위험분류

와 연관된 여러 정량적인 정보는 물론 정성적 정보에 대한 전문가의 지식을 체계적으로 도출할 수 있다. 또한 AHP는 의사결정자의 오랜 경험이나 직관 등에 기초하고 있기 때문에 계량화되어 있지 않은 정성적인 정보를 보다 쉽게 처리할 수 있는 장점을 가지고 있다(Satty, 1980). 따라서 AHP를 이용함으로써 부도예측과 관련된 정성적인 정보에 대한 전문가들의 지식(가중치)을 보다 체계적으로 도출할 수 있다.

실무 및 학계에 종사하고 있는 전문가들에 대한 설문결과에 의하면 약 80%이상의 전문가들은 정량적 정보에만 기초를 두고 있는 통계적 모형보다 전문가의 판단(Judgemental Forecasting)에 의존하여 부도가능성을 예측하는 경향이 높은 것으로 나타났다 (Bunn and Wright, 1991; Cerullo and Avila, 1975; Klein and Linneman, 1984). 부도예측에서도 전문가들의 판단에 의존한 부도예측율이 통계적모형의 부도예측율보다 더 높은 경우도 보였으며, 특히 통계적인 선형모형과 전문가 지식을 혼합한 예측방법(Combined Forecasting)을 이용할 때의 부도예측율이 통계적 모형이나 전문가의 판단에만 의존한 예측율보다 더 높게 나타났다 (Lawrence et al, 1985, 1988; Wolfe and Flores, 1990).

이러한 관점에서 본 연구에서는 전문가를 대상으로 AHP 설문 및 인터뷰를 실시하여 부도

가능성과 연관된 정량적인 정보 이외에 정성적인 정보에 대한 상대적인 중요도(가중치)를 도출하여 부도예측모형을 개발하고, 실제의 여신 자료를 이용하여 검증함으로써 전문가의 지식에 대한 특성 및 예측력을 검증하는 한편, 여러 가지 Data Mining 기법들과 비교하여 궁극적으로는 혼합예측방법의 기본방향을 설정하고자하는 연구의 목적을 두고 있다.

3.4 국·내외의 AHP기법 응용관련 선행연구

국외의 AHP기법 응용관련 선행연구는 공학에서 경영학까지 그 응용범위가 대단히 넓다. 그 중에서도 기업재무와 금융과 관련된 국외의 AHP 연구들에 대해서 살펴보면 다음 <표 4>와 같다.

<표 4> 국외 AHP기법을 이용한 기업재무 금융관련 논문

발표년도	저 자	연 구 분 야
1980	Saaty	포트폴리오 선택에 관한 연구
1980	Rogers & Pell	포트폴리오 선택에 관한 연구
1981	Saaty & Gholanmnezhad	AHP를 이용한 예측방법
1981	Wind & Douglas	포트폴리오 선택에 관한 연구
1981	Vargas & Saaty	AHP에 대한 재무적 접근
1987	Jensen	AHP에 대한 재무적 접근
1989	Vargas & Roura-Agusti	AHP를 이용한 개발도상국의 재무전략 선택
1990	Wolfe, Flores	AHP를 이용한 수익률의 예측

국내의 AHP기법 응용관련 선행연구는 다음의 <표 5>와 같이 여러 연구가 있으나 대부분의 연구가 생산관리와 경영과학분야에 치중되어 있는 것을 알 수 있다. 또한 모형개발을 중심으로 한 실증분석이 없는 연구들이 대부분이기

때문에 현실적용에는 상당한 한계점이 있어왔다. 따라서 본 연구에서는 단순히 가중치만 구하는 것이 아니라 AHP로부터 도출된 가중치를 활용한 부도예측모형을 개발하여 AHP의 실증적가치를 높이고자 하였다.

<표 5> 국내 AHP기법 응용관련 발표논문

발표년도	저 자	연 구 분 야
1991	황규승	한국의 생산자동화기술수준 예측
1993	박노국, 문희영, 송문익	AHP기법을 이용한 품질기능 평가
1994	김성천, 어하준	가중치 결정시 전문가의견종합 방법
1995	임채연	집단의사결정을 위한 AHP확장
1995	이영찬, 민재형	R&D투자계획수립에 관한 연구
1995	김영태	내부감사인 판단에 관한 연구
1996	정호원, 강인배	전자경비 시스템의 평가에 관한 연구
1996	김영대	개인정보 보호 정책의 중요도 평가에 관한 연구
1996	김형준, 김영민	전력공급원 구성비율 설정에 관한 연구
1996	변대호	자동차 구입모델 선정에 관한 연구
1996	윤재곤	MIS 성공요인 평가를 위한 3가지 통계기법의 비교
1997	우훈식, 김광용, 강성범	AHP와 LOGIT의 부도예측력 비교

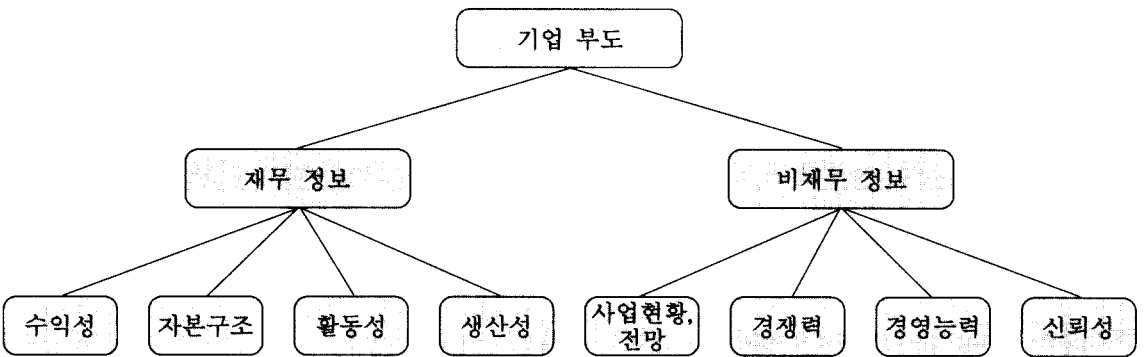
3. AHP 부도예측 모형개발

3.1 예비조사 및 계층화 구조의 설문지 작성

부도예측과 관련된 정량적인 재무정보와 정성적인 비재무정보에 대한 전문가들의 가중치를 도출하기 위한 설문지의 작성을 위하여 약 40여 일 동안에 12명의 전문가를 대상으로 한 예비 설문조사를 통하여 설문양식의 적절성을 평가하였다. 이러한 예비조사 결과 다음과 같은 결론을 도출하였다. 첫째, 대기업의 부도예측시에는 재무정보를 중요한 요인으로 간주하는 반면에 중소기업의 부도예측에서는 비재무적 정보에 비중을 더 두고 있는 것으로 나타나서 설문을 중소기업과 대기업으로 나누어 실시하여야 함을 알 수 있었다. 둘째, 부도예측 변수가 많을 수

록 전체적으로 비일관성 비율이 높아지는 문제점이 발생하였다.²⁾

따라서 본 연구에서는 AHP 설문을 대기업과 중소기업으로 나누어서 실시하였으며 또한 재무비율 15개의 변수를 4개의 요인으로, 비재무비율 11개의 변수를 4개의 요인으로 축약시켜 총 8개의 요인을 이용해 기업부도에 대한 AHP 설문을 실시하였다. 부도예측 모형의 계층구조는 [그림 1]에서 보듯이 3개의 계층으로 구성하였다. 부도가능성과 연관된 요인을 정량적인 재무정보와 정성적인 비재무정보로 구분하여 정량적인 재무정보를 다시 수익성, 자본구조, 활동성 및 생산성 요인으로 세분하는 한편 정성적인 비재무정보를 사업현황 및 전망, 경쟁력, 경영능력과 신뢰성 요인으로 세분하였다. 따라서 AHP 설문의 총 문항수는 13개였다. (우춘식 외 2인, 1997)



[그림 1] AHP에 기초한 부도예측 모형의 계층

2) 예비조사에 사용된 AHP 계층은 3개의 계층으로 총 설문 문항은 93개로 설문에 응하는 전문가는 너무 많은 설문문항에 불만을 토로하였다. Satty(19080)는 각 계층에 포함되는 속성(또는 평가기준)들의 수가 많아질수록 쌍별비교의 횟수가 증가하기 때문에 한 계층의 평가기준 수가 9개를 넘지 않는 것이 바람직하다고 하였다. 속성의 수가 너무 많은 경우에는 통계적 요인분석을 이용해서 각 계층별 중요요인을 선정할 수도 있다(Saaty, 1984)

3.2 각 요인별 변수의 선정 및 자료의 처리

본 연구에서 이용된 변수들 중 정량적 재무정보에 해당하는 변수는 한국신용평가에서 사용하는 자료를 그대로 인용하였으며, 정성적 비재무정보는 K은행의 신용조사표를 중심으로 설립년도와 자금전망의 정성적인 요인을 추가하였다.

26개의 부도예측 변수를 8개 요인으로 축약한 자료처리과정은 다음과 같다. 먼저 아래 <표 6>과 <표 7>에서 보듯이 각 요인에 해당하는 변수들을 그룹화 시킨 후에, 정량적 재무변수의 경우에는 각기 속한 산업의 평균값으로 나누어줌

으로써 산업의 특성에 의한 차이점을 극소화시켰다. 또한 각 변수들은 서로 다른 측정단위를 갖고 있으므로 그대로 합하거나 평균값을 계산하면 단위의 편의(scale bias)가 발생하므로 각 변수별로 표준화(standardization)를 하였다. 이렇게 산업평균으로 조정하고 표준화된 변수들 중 각 요인별로 해당하는 변수들의 합한 값을 계산하여 그 값이 각 요인의 대표값이 되도록 자료를 변환, 처리하였다. 이과정에서 하나의 변수라도 자료가 없으면 그 회사의 자료를 삭제한 결과 총 26개의 부도예측 변수는 4개의 재무적 요인과 4개의 비재무적 요인으로 축소되고 총 표본수는 54개로 줄어들었다.

<표 6> 정량적인 재무정보의 내용

재무요인	대표적항목
수익성요인	총 자본 순 이익률, 자기자본 이익률, 이자보상 비율, 금융비용 부담률
자본구조요인	자기자본 비율, 부채 비율, 유동 비율, 부채 상환 계수
활동성요인	총 자본 회전을, 매출채권 회전을, 재고자산 회전을
생산성요인	부가가치 및 경상이익(종업원 1인당), 노동 장비율, 자본 집약도

<표 7> 정성적인 비재무정보의 내용

비재무요인	항목	설명	측정단위
사업현황 및 전망요인	수익전망성	최근 3년간 흑자 여부	① 1년흑자 ② 2년흑자 ③ 3년이상 흑자
	판매 전망, 시장개척력	업체의 시장성과 매출액 변동액	① 시장성 불량 매출액 감소 ② 시장성 불량 매출액 정체 ③ 시장성 불량 매출액 시장 ④ 시장성 보통 매출액 감소 ⑤ 시장성 보통 매출액 정체 ⑥ 시장성 보통 매출액 신장 ⑦ 시장성 양호 매출액 감소 ⑧ 시장성 양호 매출액 정체 ⑨ 시장성 양호 매출액 신장
	업종유망성	업체의 주기상 위치	① 불안정 태동초기 ② 불안정 태동후기 ③ 안정 성숙초기 ④ 안정 성숙후기
	자금전망	자금전망에 대한 상태와 전망	① 매우 불량 ② 다소불량 ③ 균형 ④ 양호 ⑤ 매우 양호
경쟁력 요인	기업형태	기업의 형태	① 기타 ② 외감법인 ③ 등록법인 ④ 2부 상장 ⑤ 1부 상장
	기술개발, 품질혁신	기술개발과 품질혁신	① 기미달(미약) ② 평균(보통) ③ 양호(우위) ④ 훨씬양호

비재무요인	항 목	설 명	측 정 단 위			
경영능력 요인	경영자의 경영능력	현 경영자의 경영기간	①3년 미만 ⑤15년 이상	②3년 이상	③5년 이상	④10년 이상
	사력 (설립년도)	설립된 후의 기간	①90년도 이상 ④10년 이상	②88~90년도 ⑤80년 이하	③85년~88년	
	노사관계	최근의 노동쟁의 발생 현황	①최근 1년내 쟁의 없음 ③최근 3년간 쟁의 없음		②최근 2년간 쟁의 없음	
신뢰성 요인	은행거래 신뢰도	과거 은행거래시 신뢰도	①미달	②평균	③양호	④훨씬 양호
	대외신뢰도	현재 금융기관 이외의 업계평판	①미달 ③등2년간 부도연체 없음		②1년간 부도연체 없음 ④3년간 부도연체 없음	

3.3 표본기업의 선정

표본기업은 제조업체를 대상으로 하여 S 및 K보험회사의 대출기업으로서 1993년부터 1995년 동안에 부도처리된 기업 38개사와 동기간에서 무작위로 선정한 정상기업 45개사(총 83개기업) 중에서 자료이용이 가능한 기업으로 제한하여 최종표본으로 31개의 정상기업과 23개의 부도기업등 총 54개 기업을 선정하였다.

한편 부도예측 모형의 개발과 검증을 위하여 표본기업을 모형개발을 위한 분석표본(training set)과 모형검증을 위한 검증표본(test set)으로 분류하였다. 그러나 자료이용상의 제약으로 인하여 전체표본이 54개사로 제한됨에 따라, 적은 수의 전체표본을 분석표본과 검증표본으로 분류할 때 나타날 수 있는 표본분류에 의한 편향(biases)이 발생할 수가 있었다. 따라서 이러한 편향을 줄이기도 하고 각 모형의 안정성을 점검할 수 있도록 20회에 걸쳐 분석표본과 검증표본을 50:50의 비율로 무작위 추출을 반복한 후 각 표본별로 부도예측 모형의 검증을 실시하였다.

3.4 전문가의 선정

부도가능성과 연관된 여러 정량적 정보와 정성적 정보에 대한 전문가들의 지식(가중치)을 도출하기 위하여 재계의 전문가 집단으로 J, H 및 K 은행 연구소에서 각각 9명, 9명 및 12명을, H 신용평가회사에서 4명을, 학계의 전문가로 재무나 회계학 분야의 교수 중 11명을 선정하였다. 이렇게 선정된 45명의 전문가를 대상으로 1997년 2월부터 1997년 5월까지 면담 및 우편에 의한 설문조사를 실시하였다. 대학교수를 제외한 나머지 34명의 전문가들을 연령별, 학력별, 직책별, 경력기간 별로 세분화한 결과를 보면 연령별로는 30-40세가 22명으로 가장 많으며, 학력별로는 대졸이 25명, 직책별로는 기업분석실 과장이 13명, 경력기간 별로는 10년-20년이 가장 많은 것으로 나타났다.

3.5 AHP부도예측모형의 가중치 도출

<표 8>은 설문결과 각종 정보에 대한 전문가들의 가중치를 종합한 결과로 산술평균과 기하평균별로 대기업 및 중소기업의 가중치를 보여 주고 있다.

〈표 8〉 AHP로부터 추출된 전문가의 부도예측에 대한 정보의 가중치

부도예측변수	대기업가중치		중소기업가중치		
	산술 평균	기하 평균	산술 평균	기하 평균	
정량적 정보	(1) 수익성	11.75	8.20	11.07	11.32
	(2) 자본구조	14.28	9.92	10.52	9.36
	(3) 활동성	5.73	7.00	5.79	4.66
	(4) 생산성	8.33	8.20	7.41	7.96
	합 계	40.09	33.32	34.79	33.30
정성적 정보	(5) 사업현황 및 전망	18.92	18.80	18.89	16.41
	(6) 경쟁력	18.01	24.55	13.57	16.41
	(7) 경영능력	12.69	13.34	18.47	19.88
	(8) 신뢰성	10.23	10.00	62.40	14.00
	합 계	59.91	66.68	65.21	66.70

〈표 8〉에서 보여주는 것처럼 대기업과 중소기업의 가중치에서 모두 정량적인 정보보다 정성적인 정보의 가중치가 상대적으로 높게 나타나고 있다. 특히 산술평균과 기하평균 공히 중소기업의 경우가 대기업보다 정성적 정보에 더 많은 가중치를 주는 것을 알 수 있다. 한편 기하평균이 산술평균보다 정성적인 정보에 더 많은 가중치를 부여하는 것을 알 수 있다. 결론적으로 AHP모형으로부터 도출된 가중치는 정량적 정보보다 정성적 정보에서, 정성적 정보의 가치는 대기업보다 중소기업에서, 추정되는 정성적 정보의 가치는 산술평균보다 기하평균에서 더 높게 나타났다.

3.6 AHP를 이용한 부도예측모형의 개발

AHP를 통해서 산출된 전문가의 부도예측요인에 대한 상대적 가중치는 AHP를 이용한 부도예

측 모형의 개발에 이용되었다. 즉 각 요인별로 도출된 상대적 가중치를 각 요인값에 곱한 후 그들의 합을 계산하였다. 예를 들어 〈표 8〉의 기하평균을 이용한 대기업가중치의 AHP 모형의 부도예측값은 다음과 같다.

$$\begin{aligned} \text{부도예측값} &= (8.20) * \text{수익성} + (9.92) * \text{자본구조} \\ &+ (7.00) * \text{활동성} + (8.20) * \text{생산성} \\ &+ (18.80) * \text{사업전망(현황)} + (24.55) * \text{경쟁력} + \\ &(13.34) * \text{경영능력} + (10.00) * \text{신뢰성} \end{aligned}$$

이렇게 도출된 AHP 모형의 부도예측값은 분석집단의 자료를 이용하여 최적 임계치(optimal cutoff point)를 계산하여 부도기업과 건전기업으로 분류하는 부도예측모형을 개발하였다. 최적 임계치는 분석집단을 최적(가장 높은 부도예측율)으로 부도기업과 정상기업으로 분류하는 임계값을 의미한다. 이러한 방법으로 분석표본에서 유도된 최적 임계치와 AHP모형으로부터 도출된 각 변수의 가중치를 이용한 부도예측값을 검증표본에 적용하여 AHP 모형의 부도예측율을 계산하였다.

한편 AHP 모형과의 비교를 위하여 같은 방법으로 여러 가지 Data Mining 모형들을 개발하여 부도예측율을 계산하였다. 다시 말하여 20회에 걸쳐 추출된 각 분석집단을 이용해서 Data Mining 모형들을 개발한 다음에 각 분석집단에 해당하는 검증집단에서 부도예측율을 계산하여 AHP 모형과 Data Mining 모형들 사이의 부도예측율을 비교하였다.

4. 분석결과의 해석

전문가의 지식을 이용한 부도예측 모형의 개

받하기 위하여 대기업과 중소기업으로 구분하여 각 관련요인에 대한 전문가들의 가중치를 AHP를 이용하여 도출한 다음에 대기업과 중소기업 사이에 가중치가 의미있는 차이를 보이는가를 검증하기 위하여 t-검증을 하였다. 검증결과에 의하면 예비 설문조사와 달리 신뢰성 요인을 제외하고는 모든 요인에서 유의적인 차이가 없는 것으로 나타나고 있어 대기업과 중소기업에 대한 별개의 부도예측모형을 설계하는 것이 비효율적인 것으로 판명되었다. 또한 자료이용 상의 제약으로 충분한 표본기업을 확보하지 못하였을 뿐만 아니라 54개의 표본기업을 분석한 결과 표본기업 중에서 대기업³⁾이 약 80%를 차지하고 있어 중소기업의 가중치를 이용하는 것이 무의미하다고 판단하였다. 따라서 본 연구에서는 대기업의 가중치만을 이용한 AHP모형을 개발하여 검증하였다.

4.1 AHP 모형의 부도예측력

앞에서도 기술하였듯이 표본수가 너무 적은 관계로 전체표본 54개 기업을 각 50%씩 분석표본과 검증표본으로 무작위로 선정하는 과정을 20회에 걸쳐 반복하였다. 즉 무작위로 선정된 분석표본을 대상으로 AHP 모형을 설계한 다음에 검증표본을 대상으로 AHP 모형의 예측력을 측정하였다. 아래 <표 9>에 산술평균과 기하평균의 가중치를 이용하여 도출된 AHP 부도모형의 예측력을 정리하였다.

<표 9> AHP 모형의 예측력 결과

횟수	정량적 정보		정성적 정보		정량적 및 정성적 정보	
	산술 평균	기하 평균	산술 평균	기하 평균	산술 평균	기하 평균
Mean	59.72	57.36	71.57	74.20	63.46	73.70
S.D.	0.49	11.33	1.13	5.64	3.93	5.76

<표 9>에서 보여주는 것처럼 각 속성에 대한 전문가들의 가중치를 종합하는 경우, 기하평균 방법을 중심으로 하여 살펴보면 정량적 정보만을 이용한 AHP 모형의 예측율이 57.36%이고 정성적인 정보만을 이용한 AHP 모형의 예측율이 74.20%인 한편, 정량적인 정보와 정성적인 정보를 모두 이용한 AHP 모형의 예측율이 73.70%인 것으로 나타났다. 따라서 정량적인 정보만을 이용한 AHP 모형보다 정성적인 정보만을 이용한 AHP 모형의 부도예측율이 높았으며, 특히 정량적 정보와 정성적 정보를 모두 사용한 기하평균에 기초한 AHP모형의 예측율이 73.70%를 보이고 있어 AHP 부도예측모형의 경우에는 정량적 정보의 역할이 거의 없었음을 알 수 있다.

기하평균은 정량적 정보만을 이용한 경우, 정성적 정보만을 이용한 경우, 정량적 정보와 정성적 정보를 모두 이용한 경우로 분류할 때 3 가지 경우에서 모두 일관성 비율이 0.1 이하의 값을 보이고 있어 전문가들의 지식도출이 체계적으로 수집되고 종합되었다는 것을 알 수 있다. 또한 <표 9>에서 보았듯이 부도예측력도 산술평균에 비하여 높은 것을 알 수 있다. 따라서 본 연구의 분석결과에 의하면 여러 전문가의 의견을 종합하는 방법으로서 기하평

3) 중소기업과 대기업의 분류기준은 예비조사에서와 같이 종업원기준으로 300명 이상인 기업을 대기업으로 그리고 300명 미만인 기업을 중소기업으로 분류하였다.

균방법이 산술평균방법보다 우수하다고 말할 수 있다.

4.2 각 모형의 상대적 가중치

먼저 AHP 모형의 경우 기하평균법이 산술평균보다 견고성이나 예측력에서 우수하므로 이하 논문에서는 AHP 모형의 여러 전문가 의견의 종합은 기하평균을 중심으로 정리하고자 한다. 통계적 선형모형의 대표적인 기법으로 다중판별분석모형(Multiple Discriminant Analysis: DA) 모형을 이용하였고 사용된 변수의 가중치는 각 변수의 부도예측 모형의 개발시 도출된 표준화된 판별계수(standardized discriminant function coefficient)를 백분율로 환산하여 정리하였다. 인공신경망(Neural Network: NN)의 경우는 백프로파게이션 알고리즘을 사용하고 은닉층의 노드수를 여러개 실행한 결과 가장 높은 예측력을 보인 4개의 은닉노드수(hidden nodes)를 사용한 결과를 정리하였다. 특히 인공신경망의 결과를 예측하는데 기여한 변수의 상대적중요도의 지표로 사용되는 기여요인값(Contribution Factor)⁴⁾을 각 변수의 상대적 중요도의 자료로 사용하였다. ID3 모형의 경우는 If-Then rule을 형성하는데 가장 먼저 쓰인 변수에 더 많은 가중치를, 가장 나중에 쓰인 변수에 가장 낮은 가중치를 단계별로 적용한 후 백분율로 환산하여 최종적으로 합산하여 평균치를 계산하였다. 이러한 절차를 거쳐 계산된 각 모형의 상대적 가중치를 <표 10>에 정리하였다.

<표 10> 부도예측시 사용되는 변수의 상대적 가중치

부도예측 변수		AHP 모형 (기하평균, 대기업)	DA 모형	NN 모형	ID3 모형
정량적 정보	수익성	8.20	12.63	13.34	14.79
	자본구조	9.92	15.50	15.02	14.79
	활동성	7.00	14.08	7.08	5.63
	생산성	8.20	7.41	7.91	11.27
	합 계	33.32	49.73	43.35	46.48
정성적 정보	사업현황 및 전망	18.80	19.36	17.56	22.54
	경쟁력	24.55	13.71	19.90	26.76
	경영능력	13.34	12.20	10.01	0
	신뢰성	10.00	4.99	9.18	4.23
	합 계	66.68	50.27	56.65	53.52

위의 <표 10>에서 보듯이 각 모형은 정량적 정보보다 정성적 정보에 더 많은 가중치를 주고 있음을 알 수 있으며, 특히 전문가의 판단은 Data Mining 기법들에 비하여 훨씬 더 높은 가중치를 정성적 정보에 주고 있음을 알 수 있다. 또한 여러 가지 Data Mining 기법중 NN이 전문가의 가중치와 가장 근접한 패턴을 보이는 것을 알 수 있다. 즉 정성적 정보에 대한 가중치도 전문가의 판단에 가장 근접하여 있으며 상대적 가중치의 순위나 정도도 전문가의 판단에 가장 근접하여 있다는 것을 알 수 있다. 물론 이러한 결론을 일반화 시켜서 이론화하기에는 현재까지 미흡한 면이 있으나 일단 각 모형에서 사용되는 변수의 상대적인 중요도를 선택적으로 비교 연구한 것은 향후 각 모형의 장단점을 파악하여 예측력의 원천(source)을 밝히거나 전문가의 지

4) 동일한 네트워크에서 출력변수를 예측할 때, 다른 입력변수와 비교하여 기여한 상대적 중요도를 말하며 이것은 네트워크학습시 도출된 Weight로부터 계산된다.

식과 Data Mining으로부터 도출된 지식과의 호환성(comparability)을 밝혀 IDSS(intelligent decision support system)를 설계하고 구축하는데 많은 도움을 줄 것으로 기대된다.

4.3 각 모형의 예측력

4.3.1 다변량판별분석 모형

〈표 11〉 다변량 판별분석의 결과

횃수	정량적 변수		정성적 변수		정량·정성적 변수	
	분석 표본	검증 표본	분석 표본	검증 표본	분석 표본	검증 표본
Mean	78.32	63.50	80.94	74.35	86.70	70.32
S.D.	6.43	7.11	8.51	6.84	8.50	8.98

〈표 11〉에서 보는 것과 같이 DA를 이용한 부도예측모형은 정성적자료만을 이용한 모형이 74.35%로 가장 높으며, 그 다음이 정량 정성적 자료를 이용한 모형(70.32%)이고, 정량적 자료만을 이용한 모형(63.50%)은 가장 낮은 결과를 나타내고 있다. 또한 표준편차에서도 정량적 자료(7.11), 정량 정성적 자료(8.98)보다 정성적 자료(6.84)를 이용한 모형의 편차가 적은 것으로 나타나 정성적 정보를 이용한 모형이 더 안정적(robustness)임을 나타내준다. 위의 결과로 볼 때 DA를 이용한 기업도산예측 모형의 개발에 있어 정량적인 정보뿐만 아니라 정성적인 정보를 고려해야 한다고 할 수 있다.

4.3.2 인공신경망모형

인공신경망모형은 학습알고리즘, 은닉층의 수, 학습의 횃수에 따라 예측결과의 차이가 있어 본

모형에서는 백프로과제이션 학습알고리즘을 사용하고, 학습의 횃수는 1,000-10,000을 중심으로 500번이상 오류 극복이 없으면 중단하였다. 특히 은닉층의 수에 따라 판별력은 상당한 차이를 보이므로 최적 은닉층 노드수의 결정을 위하여 여러 가지 수를 점검하였다. 본 연구의 경우 은닉층이 4개일 때 가장 좋은 예측력을 나타냈으므로 최종신경망 모형의 은닉층 노드수는 4개로 지정하였다.

〈표 12〉 인공신경망을 이용한 도산예측 결과

횃수	정량적 변수		정성적 변수		정량·정성적 변수	
	분석 표본	검증 표본	분석 표본	검증 표본	분석 표본	검증 표본
Mean	87.05	64.62	93.35	75.23	94.10	74.57
S.D.	9.38	11.03	4.23	7.63	7.63	10.04

〈표 12〉에서 보는 것과 같이 NN 모형은 정성적 모형(75.23%)과 정량 정성적 모형(74.57%)이 정량적 모형(64.62%) 보다 아주 높은 예측력을 보이고 있음을 알 수 있다. 또한 표준편차는 정성적 모형(7.63)이 정량적 모형의 (11.03)보다 낮아 정성적 모형이 좀더 안정적이라 할 수 있다. 따라서 NN을 이용한 부도예측모형의 설계시 DA와 마찬가지로 정량적 변수뿐만 아니라 정성적 변수를 고려해야함을 시사해준다고 할 수 있다.

4.3.3 ID3 모형

〈표 13〉은 ID3를 실행시킨 결과이다. 여기서 주목하여야 하는 것은 ID3는 사용된 정보가 매우 깨끗하게 위험분류가 가능하여야 한다는 것이다. 〈표 13〉에서 보듯이 ■은 사용된 정보의 불충분으로 인하여 분석표본에서 If-Then Rules

를 생성하지 못한 경우이고, □은 사용된 정보의 불완전성으로 인한 경우, 즉 하나의 데이터셋에서 서로 상반(conflict)되는 데이터가 존재할 경우에 일부 속성이 no-data로 처리되어 해당 속성은 If-Then Rules에 제외되는 경우에 해당된다.

〈표 13〉 ID3를 이용한 도산예측 모형의 결과

ID3		정량	정성	정량,정성
Set 1		68.4 □	60.0	60.0
2		63.3	82.1 □	60.7 □
3		85.2	■	85.1
4		66.7	66.7 □	95.5 □
5		58.3	58.3	66.7
6		57.1	74.3	60.0
7		51.8	74.1	51.8
8		89.5	78.9	89.5
9		72.0	92.0	56.0
10		100.0	■	100.0
11		66.7	■	65.4 □
12		64.0	80.0 □	■
13		46.7	88.9 □	■
14		75.0	100.0□	92.3 □
15		57.1	78.6	82.1
16		80.0	95.3 □	68.0 □
17		48.2	85.2	92.0 □
18		■	78.3	100.0□
19		74.7	74.7 □	80.8
20		69.7	87.1 □	72.7
ID3__I	Mean	68.1	80.4	76.6
	S.D.	14.0	11.0	16.0
ID3__II	Mean	61.3	62.1	62.0
	S.D.	11.0	10.0	12.0

따라서 본 연구에서는 ■나 □가 존재하는 경우, 검증표본에서 분류되지 않은 회사의 부도 위험을 부도 예측시 삭제시킨 경우를 ID3_I으로 명명하여 정리하였고, 검증표본에서 분류 못한 경우도 잘못된 예측으로 고려하여 부도예측력을 계산한 경우를 ID3_II로 명명하여 별도로 정리하였다. 본 연구시 발생한 이러한 현상과 관련해서 유추하여 보건대 기존 연구 중에서 ID3를 포함한 모형간의 예측력비교는 한번쯤 재고 될 필요가 있는 것으로 사료된다. 즉 기존의 ID3 예측력은 본 연구와 같은 현상이 일어났을 경우 부도예측이 불가능한 자료에 대하여 어떤 처리를 하였는지 재점검이 필요할 것으로 보인다.

ID3는 분석표본의 예측력이 100%이므로 분석표본의 결과는 생략하였다. 〈표 13〉에서 보는 것과 같이 두 모형 모두 정성적 모형의 예측력이 정량적 모형보다 높게 나타나며, 표준편차 역시 정성적 모형이 가장 낮은 것으로 나타났다.

4.3.4 각 모형의 예측력 및 안정성 비교

앞에서 기술한 여러 가지 모형의 예측력과 안정성을 〈표 14〉에 비교, 정리하였다. 정량적 정보만을 이용한 모형의 경우에는 평균부도예측력은 NN모형이 우수하고 안정성 측면에서는 DA모형이 약간 우수한 것으로 나타났다. 정성적 정보만을 이용한 경우에는 평균부도예측력에서는 NN이 약간 우수하고 안정성 면에서는 AHP모형이 약간 우수한 것으로 나타났다. 또한 정량 정성적 변수 모두를 이용한 모형에서는 예측력에서 NN이 약간 우수하고, 안정성에서는 AHP모형이 우수한 것으로 나타났다. 따라서 예측력 측면에서는 NN이 사용된 정보에 관계없이 전반적으로 가장 우수하였으며, 안정성을 고려할 경우 정량적 정보의 경우는 DA모형이 우수

하고, 정성적 정보나 또는 모든 정보를 사용한 경우에는 AHP 모형이 우수한 것으로 나타났다.

〈표 14〉 각 모형의 부도예측력

부도예측변수	평균부도예측율/표준편차			
	AHP기 하평균	DA	NN	ID3
(1) 정량적 정보	57.36% 11.33	63.50% 7.11	64.62% 11.03	61.32% 11.00
(2) 정성적 정보	74.20% 5.64	74.35% 6.84	75.23% 7.63	62.08% 10.01
(3) 정량적 및 정성적 정보	73.70% 5.76	70.32% 8.98	74.57% 10.04	62.00% 12.03

ID3의 경우를 보면 의사결정트리를 유도하는데 있어 AHP, DA, NN 에 비해 사용된 변수는 적음에도 불구하고 정량적 자료를 이용한 경우에는 다른 모형에 비하여 예측정확성에 큰 차이

가 나지 않고 있다. 그러나 정성적 모형과 정량 정성적 모형에서는 다른 모형에 비하여 부도예측정확성이 많이 떨어지는 것으로 나타났다. 또한 ID3의 경우 분석에 사용된 정보의 유형에 관계없이 일정한 예측정확성을 나타내고 있어 정보의 유형에 따라 안정적임을 보여주고 있다. 반면에 표준편차는 보편적으로 가장 높은 것으로 보아 ID3의 단점인 투입된 변수에 최적화하는 경향이 있어 견고성(robustness)이 떨어진다는 것을 알 수 있다. 즉 기존에 생성된 If-Then Rule들이 조그마한 변화에 매우 민감하게 반응하여 If-Then Rules 자체를 180도 바꾸는 현상이 나타났다.

이러한 각 모형별 부도예측력이 통계적으로 유의한 차이가 있는지 알아보기 위하여 사용된 정보의 특성별로 일원분산분석을 실시하고 그 결과를 〈표 15〉에 정리하였다.

〈표 15〉 사용된 정보별로의 모형의 부도예측력에 대한 ANOVA 결과

부도예측변수	Source	Degree of Freedom	Mean Square	F Ratio	p value
(1) 정량적 정보	Between Groups	3	204.70	1.6631	.1822
	Within Groups	75	123.08		
(2) 정성적 정보	Between Groups	3	695.72	10.8424	.0000
	Within Groups	73	64.17		
(3) 정량적 및 정성적 정보	Between Groups	3	612.49	5.3618	.0021
	Within Groups	74	114.23		

ANOVA 분석결과 정량적 정보의 경우 모형별로 부도예측력의 차이는 없는 것으로 나타났으나 정성적 정보나 모든 정보의 사용시에는 모형별로 부도예측력이 차이가 있는 것으로 나타났다. 따라서 어느 모형의 부도예측력이 차이가 있는지를 확인하기 위하여 Tukey의 LSD를 이용하여 사후다중비교(Multiple Comparison Test)를 하였다. 그 결과 정성적 정보의 경우에는

ID3와 그외 다른 모형들(DA, NN, AHP)의 예측력이 차이가 있는 것으로 밝혀졌다. 즉 DA, NN, AHP 모형의 부도예측력은 차이가 없으나, 이들 모형의 부도예측력은 ID3 모형의 부도예측력보다 우수한 것으로 판명되었다. 또한 모든 정보를 이용하는 경우에도 마찬가지로 DA, NN, AHP 모형의 부도예측력은 ID3 모형의 부도예측력보다 보다 우수하여 통계적으로 유의

한 차이가 있는 것으로 밝혀졌다.

따라서 본 연구의 특징인 전문가의 지식을 추출한 AHP 모형이 사용된 정보의 특성에 관계 없이 부도 예측력면에서 다른 Data Mining 기법보다 동등하거나 우수한 결과와 정성적 정보를 사용한 경우에는 안정성 측면에서 보다 우수하게 나타난 결과는 주목할 만하다. 비록 선형관계를 중심으로 전문가의 지식을 추출하고 모형을 개발하였다 하더라도 높은 예측력과 안정성을 보인 것은 향후 AHP 방법이 단순히 사용된 변수의 가중치만 도출하는 것이 아니라 실용화된 위험예측모형으로 발전되어 사용될 수 있다는 것을 의미한다. 특히 AHP 모형은 정성적 정보의 처리가 쉽고 부도예측력과 안정성에서도 우수한 결과를 보이므로 향후 계량화가 어렵고 신뢰성이 떨어지는 정성적 정보의 처리에 상당한 기여를 할 것으로 기대된다. 또한 Bunn & Wright(1991)가 주장하듯이 인간판단과 계량적 모형과의 혼합된 방법이 향후 불확실한 상황에 대한 예측에 있어 새로운 패러다임이 될 것이라는 관점에서 볼 때, 본 연구는 전문가의 지식과 Data Mining 지식을 통합하는 모형의 개발에 실증적 근거자료를 제시하고 있다고 본다.

마지막으로 본 연구의 분석결과, 전문가의 지식이나 Data Mining 모형 모두 다 정성적 정보에 보다 높은 가중치를 주고 있음을 알 수 있었다. 또한 여러 부도예측모형에서 전반적으로 정성적 정보만을 이용한 경우가 가장 높은 예측력을 보이고 있음을 볼 수 있었다. 따라서 이러한 결과를 종합하여 볼 때 정성적 정보의 사용이 부도예측에 있어 매우 중요하다는 결론을 내릴 수 있다. 특히 재미있는 결과는 모든 정보를 이용할 경우 모든 모형에서 부도예측력이 정성적 정보만을 이용하는 경우보다 예측력이 떨어지고 있다는 것이다. 이러한 결과로부터 정량적

정보의 가치가 없다고 결론을 내리기에에는 무리가 있으나, 확실한 것은 부도예측시 정성적 정보의 가치는 매우 크다는 사실은 확인되었다.

5. 결 론

부도가능성을 예측하는데 이용되는 대부분의 통계모형이 정량적인 정보에 의존하고 있어 정성적인 정보의 중요성을 간과하고 있다. 이러한 관점에서 본 연구는 정량적인 정보에만 의존하는 경우와 정성적인 정보에만 의존하는 경우에서 부도예측력 사이에 어떤 차이가 있으며, 부도예측에서 정성적인 정보의 중요성을 검증하고자 하였다. 특히 전문가 집단을 대상으로 한 AHP 설문조사를 실시하여 각 정보별 상대적 가중치를 도출한 후, AHP 부도예측모형을 설계하여, 이를 여러 가지 Data Mining 모형들과 비교, 분석하였다. 검증결과에 의하면 AHP 모형이 부도예측력에서 여러 다른 Data Mining 모형과 동등하거나 우수하였으며, 특히 정성적 정보를 사용하는 경우 안정성도 뛰어난 것으로 판명되어 향후 AHP 모형이 부도예측모형으로 발전될 수 있다는 것을 실증적으로 검증하였다. 특히 AHP 모형뿐만 아니라 DA, NN, ID3등 여러 Data Mining 모형에서도 정량적 정보보다 정성적인 정보가 더 중요한 역할을 한다는 증거가 발견되어 위험예측시 정성적 정보가 중요하게 작용한다는 Bunn & Wright(1991)의 연구와 동일한 결과를 얻을 수 있었다.

그러나 본 연구는 다음과 같은 점에서 한계점을 지니고 있다. 부도기업의 자료이용 상의 제약으로 인하여 부도예측과 관련된 다양한 정성적 정보를 수집할 수 없어서 표본기업을 폭넓게

선정하지 못하였다. 54개 표본기업 중에서 중소기업이 불과 21% 정도에 지나지 않아 중소기업의 가중치에 기초한 AHP 모형을 설계하여 검증할 수가 없었다. 이 점에 대하여는 향후의 연구에서 보완하고자 한다. 한편 전체표본이 너무 작은 관계로 20회에 걸쳐 전체표본을 통제표본과 검증표본으로 무작위 추출하여 분류함으로써 표본분류에 의한 편의(biases)를 통제하고자 하였으나 많은 기업으로 표본을 구성하는 것이 보다 바람직할 것이다.

참고문헌

- 김광용, "보험사기 적발을 위한 전문가시스템의 개발: 퍼지이론과 AHP를 중심으로", 보험개발연구, 제18호, pp4-28, 1996
- 김광용, "Knowledge Discovery from the hybrid Model combined by computer power and human reason", '96 추계전문가시스템 학술회의, pp325-332, 1996
- 김광용, "Expert's Reliance for the knowledge from Data Mining Techniques", working paper, 1997
- 김선기·황석하·김건우, "기업부실화의 원인, 징후 및 예측", 한국신용평가(주), 1986
- 김성천·어하준, "AHP 가중치 결정에서의 전문가 의견종합 방법", 한국경영과학회지, 1994, 12.
- 김영태, "사원부정의 발생원인과 방지방법의 중요성에 대한 내부감사인 판단에 관한 연구", 회계학 연구, 1995, 12.
- 김형준·김영민, "AHP를 이용한 전력공급원 구성비율 설정에 관한 연구", 한국경영과학회지, 1996, 10.
- 박노국·문희영·송문익, "품질기능 전개와 AHP 기법을 이용한 기능평가", 한국품질관리학회지, 1993, 12.
- 변대호, "AHP를 이용한 자동차 구입모델선정에 관한 연구", 한국경영과학회지, 1996, 10.
- 손정식·강병호·허영빈, 금융기관경영론, 박영사, 1990
- 우춘식·김광용·강성범, "LOGIT 분석과 AHP 분석을 이용한 부도예측모형의 비교연구", 재무관리연구, 14(2), 229-252, 1997
- 이영찬·민재형, "불확실한 상황 하에서 다목표 R&D 투자계획수립에 관한 연구", 한국경영과학회지, 1995, 8.
- 이재식·한재홍, "인공신경망을 이용한 중소기업 도산예측에 있어서의 비재무정보의 유용성검증", 한국전문가시스템학회지, 1(1), 1995
- 이재식·한재홍, "인공신경망모형과 사례기반추론 모형의 기업도산 예측률 비교", '96 전문가시스템학회 추계학술대회, 1996
- 윤재근, "MIS성공요인평가를 위한 3가지 통계기법의 비교", 한국경영과학회지, 1996, 12.
- 정호원·강인배, "AHP를 이용한 전자경비 시스템의 평가에 관한 연구", 한국경영과학회지, 1996, 6.
- 황규승, "한국의 생산자동화기술수준 예측", 한국경영과학회, 1991, 12.
- Altman, E., "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance*(September), pp.589-609, 1968.
- Bunn, D. & Wright, G., "Interaction of Judgemental and Statistical Forecasting Methods: Issues & Analysis", *Management Science*, 1991.

- Cerullo, M. J. & Avila, A., "Sales Forecasting Practices: A Survey", *Managerial Planning*, 24, pp.33-39, 1975.
- Edmundson, R.H., Lawrence, M.J., & O'Connor, M.J., "The Use of Non-Time Series Information in Sales Forecasting: A Case Study", *Journal of Forecasting*, 7(3), pp.201-212, 1988.
- Foster, G., *Financial Statement Analysis*. 2nd ed. Prentice-Hall, Chapter 15. 1986.
- Jensen, R. E., "A Dynamic Analytic Hierarchy Process Analysis of Capital Budgeting Under Stochastic Inflation Rates, Risk Premiums, Liquidity Preference: Theory", *Advances in Financial Planning and Forecasting* 2, pp.269-302, 1987a.
- Jensen, R.E., "International Investment Risk Analysis: Extensions for Multinational Corporation Capital Budgeting", *Mathematical Modeling*, 9, pp.265-284, 1987b.
- Klein, H. E. & Linneman, R.E., "Environmental Assessment: An International Study of Corporate Practice", *Journal of Business Strategy*, pp.66-84, 1984.
- Lawrence, M.J., Edmundson, R.H., & O'Connor, M.J., "An Examination of the Accuracy of Judgemental Extrapolation of Time Series", *Journal of Forecasting*, pp.14-25, 1985.
- Marais, M., Patell, J., & Wolfson, M., "The Experimental Design of Classification Models: An Application of Recurse Partitioning and Bootstrapping to Commercial Bank Loan Classification", *Journal of Accounting Research* 22(Supplement), 1984.
- Rothe, J. T., "Effectiveness of Sales Forecasting Methods." *Industrial Marketing Management*, pp.114-118. 1978.
- Rust, R.T. & Schmittlen, D.C., "A Bayesian Cross-Validatized Likelihood Method for Comparing Alternative Specification of Quantitive Models" *Managerial Science*, 4(1), pp.21-40, 1985.
- Saaty, T.L., *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980, 1984.
- Saaty, T.L. & Gholammehad, H., "High-Level Nuclear Wastement: Analysis of Options", *Environment and Planning*, B9, pp.181-196, 1982.
- Vargas, L.G., & Saaty, T.L., "Financial and Intangible Fctors in Fleet Lease or Buy Decisions", *Industrial Marketing Management* 10, pp.1-10, 1981.
- Vargas, L.G., & Roura-Agusti, J.B. "Business Strategy Formulation for a Financial Institution in a Developing Country". in: B.L. Golden, E.A. Wasil and P.T. Harker(eds) *Applications of the Analytic Hierarchy Process*", SpringerVerlag, New York, 1989.
- Wolfe, C. and Flores, B., "Judgemental Adjustment of Earnings Forecasting", 9, 4, pp.389-406, 1990.
- Zahedi, F., "The AHP: A survey of the Method and its Applications", *Interface*, 16(4), pp.96-108, 1986.
- Zavgren, C., "The Prediction of Corporate failure: The state of the art", *Journal of Accounting Literature* 2, pp.1-37, 1983.