

여러가지 Data Mining 기법으로부터 도출된 지식에 관한 전문가의 신뢰도에 대한 실증적 연구*

김광용*

An Empirical Study for the Expert's Reliance on the Knowledge from the Several Data Mining Techniques*

GwangYong Kim*

요 약

본 연구는 여러 가지 데이터마이닝 기법으로부터 도출된 지식이 어떻게 인간의 판단에 영향을 미치는가를 지식구조별, 자료특성별, 전문가지식의 일관성별로 실증적 연구를 하여 궁극적으로 전문가 의사결정에 도움이 되는 데이터마이닝 기법의 활용방안을 제시하고자 한다. 분석결과 전문가들의 판단은 데이터마이닝의 지식표현형태에 의한 영향을 많이 받고 있는 것으로 나타났으며, 특히 IF-THEN의 형태로 표현되는 명제형 지식구조에 가장 많은 신뢰를 갖고 있는 것으로 나타났다. 특히 자료의 특성, 또는 전문가의 판단 일관성과 데이터마이닝 기법 사이에 상호작용효과가 있어 향후 데이터마이닝 기법을 활용하여 전문가의 의사결정을 돕고자 할 때는 이러한 차이점을 고려해야 하는 것으로 밝혀졌다.

주제어: 데이터마이닝, 전문가의 신뢰도, AHP, 부도예측

* 이 논문은 1997년 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음

* 숭실대학교 경영학부 교수

1. 서론

본 연구의 목적은 여러 가지 데이터마이닝(Data Mining) 기법들로부터 도출된 지식에 대한 전문가의 신뢰도를 실증적으로 연구하고자 한다. 즉 귀납적 학습방법(Rule Induction), 인공신경망(Artificial Neural Network), 통계적 선형모형(Statistical Linear Model)과 같은 여러 가지 데이터마이닝 기법들이 도출한 지식들에 따라 전문가의 판단이 어떤 영향을 받는가에 대한 실증적 연구가 본 연구의 주된 목적이다.

과거의 많은 연구는 위험을 분류하고 예측하는데 있어 인간의 판단(Human Judgement)과 통계적 선형모형 중 어느 것이 우수하느냐에 그 초점을 맞추어왔다. 또한 근래에는 인공지능(Artificial Intelligence)에 기초한 많은 데이터마이닝 기법들이 통계적 선형모형과 비교하여 볼 때 예측을 측면에서 동등하거나 우수하다는 것을 보여주어 그 유효성을 인정받고 있다(Langley & Simon, 1995, Woong et al, 1997, Woong & Selvi, 1998, Zopounidis, 1997). 그러나 Hart(1992)가 지적했듯이 기계학습(Machine Learning) 기법으로부터 도출된 지식을 전문가의 검증없이 그대로 사용하는 것은 때때로 매우 위험하며, 이러한 지식들을 최종사용자가 어떻게 받아들이느냐는 것은 데이터마이닝 기법의 실제 활용측면에서 매우 중요한 문제라 할 수 있다. 그러나 현재까지 인공지능과 통계학적 방법을 포함한 데이터마이닝 기법과 전문가 판단과의 체계적인 상호작용(Structured Interaction)에 대한 연구는 많이 이루어지지 않고 있다 (Bunn & Wright, 1991).

따라서 본 연구는 여러 가지 데이터마이닝 기법으로부터 도출된 지식에 관한 전문가의 신뢰도를 (1)지식표현형태별(명제형 지식구조, 반명

제형지식구조, 비명제형지식구조(Propositional Knowledge, Pseudo-propositional Knowledge, and Non-propositional Knowledge)), (2)자료특성별(정량적 대 정성적(Numeric vs Non-numeric)), (3)전문가의 판단 일관성(Judgement Consistency)별로 분류한 실증적 연구를 하여 궁극적으로 전문가의 의사결정에 도움이 되는 데이터마이닝 기법의 활용 방안을 제시하고자 하였다.

연구결과 전문가들은 데이터마이닝의 지식표현형태에 의한 영향을 많이 받고 있는 것으로 나타났으며 특히 IF_THEN의 형태로 표현되는 명제형 지식구조에 가장 많은 신뢰를 갖고 있는 것으로 나타났다. 특히 자료의 특성이나 전문가의 판단 일관성과 데이터마이닝 기법 사이에 상호작용효과가 있어 향후 데이터마이닝 기법을 활용하여 전문가의 의사결정을 돕고자 할 때는 이러한 차이점을 고려해야 하는 것으로 밝혀졌다.

본 연구는 2장에서 전문가의 판단에 관한 문헌연구를, 제3장에서 연구방법을 제시하였으며, 제4장에서는 연구결과를 분석하였다. 마지막으로 제5장에서 결론과 향후 연구방향을 제시하였다.

2. 문헌연구

2.1 인간의 의사결정

인간의 의사결정과 관련된 연구는 크게 규범론(Normative Theory), 기술론(Descriptive Theory), 인도론(Prescriptive Theory)의 3가지로 나누어 볼 수 있다(Bell et al, 1991; Zey, 1992). 규범론은 여러 가지 공리에 기초한 엄격한 가정하에서 연역적(Deductive)인 논리에 의하여 주어진 상황에서는 반드시 이렇게 "해야만 한다(ought to)"는 결론을 갖게 한다. 대표적인 예로는 기존 여러 학

문분야에서 광범위하게 사용되는 주관적 기대효용 이론(SEU: Subjective Expected Utility)을 들 수 있다.

1970년대 초부터 주로 사회과학을 연구하는 학자들을 중심으로 기존의 규범론을 비판하는 실증적 연구가 많이 발표되었는데 이들은 “주어진 의사결정 상황에서 실제로 사람들은 어떻게 행동하고 있는가?”에 연구의 초점을 두었다. 이들은 여러 가지 실험을 통하여, 실제 많은 경우 사람들은 규범론적인 의사결정을 하지 않는다는 것을 귀납적(Deductive)인 논리에 의하여 증명하고, 제반 경영환경에서 내리는 사람들의 의사결정 모형을 제시하였다. 대표적인 예로는 이득과 손실에서 서로 다른 위험선호도를 보이는 사람의 의사결정을 이론화한 프로스펙 이론(Prospect Theory)을 들 수 있다(Tversky & Kahneman, 1981).

이렇게 규범론과 기술론이 이론의 우위성을 주장하고 있을 때, 최근에는 이 두 학풍의 중요성을 동시에 인정하면서 양자의 보완적 발전을 돕고자 하는 새로운 접근법, 즉 인도론(Prescriptive)이 등장하였다. 인도론은 “이론과 실제에서 차이를 보이고 있는 사람의 의사결정을 어떻게 도울 것인가?”에 그 초점을 두고 실제 사람들의 의사결정시에 보이는 행태적 측면의 실재를 인정하고 여러 가지 과학적이고 체계적인 방법에 의하여 합리적인 의사결정을 할 수 있도록 돕자고 하는 것이 연구의 주제이다. 특히 전문가의 판단에 대한 이러한 인도론은 전문가 판단에 관한 연구의 초점이 실험실 연구에서 실제 현장에서 일하고 있는 전문가로 옮겨가면서 실제 전문가의 지식을 중심으로 연구하자는 생태적 유효성(Ecologically Vaild)이 강조되었다(Bolger & Wright, 1994). 따라서 이러한 연구의 흐름은 전문가 판단에 대한 연구가 심리적 측면의 연구에서 전문

가 지식의 생성, 축적, 응용 과정에 대한 연구로 옮겨지면서 여러가지 인공지능기법을 활용한 전문가 지식의 연구에 대한 토대를 마련하였다고 할 수 있다(Bell et al., 1991; Langley and Simon, 1995; Lian, 1992; Turban & Aronson, 1998; Zey, 1992)

2.2 전문가의 판단과 통계적 선형모형의 위험예측력 비교

인간의 의사결정에 있어 통계적 선형모형의 유용성(Dawes & Corrigan, 1974)과 그와는 반대로 인간 판단에 있어 통계적 선형모형의 부적절성(Zeleny, 1976)이 논의된 이후, 전문가의 판단과 통계적 선형모형의 위험예측력비교에 대한 연구는 여러 분야에서 매우 활발히 진행되었다.

특히 과거에는 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 Judgement)과 통계적 선형모형과의 비교가 많이 이루어져왔다(Bunn & Wright, 1991). 일반적으로 많은 연구에서 통계적 선형모형이 인간판단보다 위험예측력면에서 우수하거나 동등하였는데, 인간의 판단이 우수한 경우만을 제시하면 [표 1]과 같다(김광용 외 2인, 1988).

[표 1]에서 보듯이 인간의 판단이 우수한 경우는 통계적 선형모형의 도움없이 비선형적 추가정보를 이용한 경우나 또는 통계적 선형모형의 도움을 받은 경우는 선형적, 정량적 자료를 이용한 경우에서 인간의 판단이 우수함을 알 수 있다. 또한 현재까지 이루어진 많은 연구는 인간의 총체적 판단(Holistic Judgement)과 여러 가

(표 1) 인간의 판단과 통계적선형모형의 실증적 비교

	인간판단	실험대상	적용문제	특 징
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	추가정보의 사용시 인간판단우수
Casey & Selling (86)	Seperate	Expert	Financial Distress	추가정보의 사용시 인간판단우수
Reinmuth & Guerts (72)	Adjusted	Expert	Sales Forecast	통계적선형모형의 이용시 인간판단 우수
Mathews & Diamantopoulos (89)	Adjusted	Expert	Sales Forecast	통계적선형모형의 이용시 인간판단 우수
Huss (86)	Adjusted	Expert	Sales Forecast	통계적선형모형의 이용시 인간판단 우수
Wolfe and Flores (90)	Adjusted	Student	Forecasting Earning	통계적선형모형의 이용시 인간판단 우수

지 통계적 선형모형과의 비교가 주종을 이루었다(Arizne & Narasimha, 1994; Chung and Silver, 1992; Hargaford & Breiman, 1983; Liang et al, 1990; Shaw, 1993)

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 우선되어야한다(Bunn & Wright, 1991). 특히 근래에는 통계적 선형모형 이외에도 인공신경망과 같은 인공지능 기법을 이용한 많은 데이터마이닝 기법들이 사용되고 있어, 이들의 사용이 인간의 판단에 어떤 영향을 주는지에 대한 연구도 같이 병행되어야 할 것이다 (Katten et al., 1993; Kim et al., 1997).

2.3 데이터마이닝 기법으로부터 도출된 지식의 특징

데이터마이닝의 기법으로 들 수 있는 것은 여러 가지가 있으나 본 연구에서는 데이터마이닝의 대표적 기법으로 귀납적학습방법(Rule Induction), 인공신경망, 통계적 선형모형 등 세 가지만을 예로 사용하였다. 아래의 [표 2]에서 보듯이 각각의 방법들은 각기 다른 특징을 가지고 있으며 특히 인지적 측면(Epistemology)에서 보면 각 방법의 지식구조(Knowledge Structure)는 상이한 특징을 가지고 있다(Gim, 1995; 김광용외 2인, 1998).

먼저 귀납적 학습방법은 "IF-THEN"의 형태로 표현되는 규칙(Rule)들의 집합형태인 명제형 지식(Propositional Knowledge)이다(Quinlan, 1986). 이러한 명제형 지식 구조는 쉬운 이해력과 분석력(Perspiciuity)을 가지고 있다. 따라서 전문가는 이러한 지식구조로부터 지식을 얻거나 처리할 때 특별한 기호학적 처리가 불필요하다(Newell & Simon, 1972). 이러한 지식은 불명확하거나 애매 모호하지 않고 간단 명료하면서도 매우 논리적이다. 따라서 전문가가 이러한 지식을 받아

[표 2] 세가지 데이터 마이닝 기법의 특징

	귀납적학습방법	통계학적 접근	NN
본 논문에서 사용된 방법	CHAID	다중판별분석(MDA)	인공신경망(NN)
최적화 기준	Chi-Square	공분산행렬 (Covariance Matrix)	일반화된 델타규칙 (Generalized Delta Rule)
최적화 절차	반복적인 분할 (Repetitive Decomposition)	행렬을 이용한 통계적 추정 (Matrix Operation)	역전파학습 (Back Propagation)
모델의 가정	분할을 위한 충분한 자료	여러 가지 통계적 가정	없음
지식표현형태	IF-THEN 규칙 (의사결정나무)	선형결합 모형 ($Y = \sum A_i X_i$)	가중치 네트워크 구조
지식의 구조	명제형 지식	반명제형지식	비명제형지식
주요장점	설명력, 쉬운이해성 (perspicuity)	설명력, 검증성 통계적 유의수준 제공	안정성(Robustness) 예측력(Predictability)
주요단점	상충되는 예의 처리어려움	가정의 만족이 어려움	설명력의 부재 (Black Box)

들이거나, 비평하거나, 또는 수정보완 할 경우 매우 쉽게 할 수 있다.

통계적 선형모형은 반명제형 지식(Pseudo-propositional Knowledge) 구조를 가지고 있다. 즉 사용된 각 변수의 상대적 가중치와 그에 대응하는 변수의 값을 곱한 후, 그 결과들의 합에 기인해서 의사결정을 하는 것은 명제형 지식(Propositional Knowledge)구조이지만 각 변수의 통계적 추정은 비명제형 지식(Non-propositional Knowledge) 구조를 가지고 있다. 이러한 비명제형 지식구조는 귀납적 학습방법(Rule Induction)에 비해서 상대적으로 낮은 이해력과 분석력을 가지고 있다. 따라서 전문가는 이러한 지식을 해석하거나 분석하는데 있어 어려움을 겪을 것이며, 받아들이나 거부하느냐의 결정도 매우 어려울 것이다. 과거 많은 연구들도 만일 사용되는 변수가 많을 경우 인간은 명제형 지식구조를 선호한다는 결과를 보여주었다 (Brehmer, 1987; Hagaford & Breimen, 1983; Olshavsky, 1979; Paquette & Kida, 1988; Payne, 1976; Rothstein,

1986).

마지막으로 NN은 지식의 집적화(Coding)을 추론(inference)보다는 패턴인식(Pattern Recognition)에 의해서 한다(Rumelhart et al., 1986). 즉 지식의 표현이 각 층(layer)별 노드(node)사이의 연결된 선의 가중치로 표현되는 비명제형 지식(Non-propositional Knowledge) 구조를 갖고 있다. 따라서 전문가가 이러한 지식을 해석하거나 분석하는 것은 거의 불가능하다 (일명 'Black Box'). 그러나 NN의 강점인 안정성(Robustness) 및 학습력(Learning Ability), 예측력(Prediction Power)은 이미 여러 연구에서 입증되었으므로 이러한 장점을 그대로 가지면서 지식의 표현은 명제형 지식구조를 갖게 하는 혼성모형이 바람직하다(Whalen & Gim, 1996).

2.4 본 연구와 관련된 기존 연구의 주제들

Payne(1976)은 인간이 데이터를 이용하여 의사결정을 하고자 할 때 다음과 같은 3가지 전략을 사용한다고 하였다. 첫째, Compensatory 전략

으로 먼저 결과(Outcome)와 관련된 중요한 변수(Attributes)를 선정하고, 각 변수에 가중치(Weight)를 부여하여 의사결정을 하는 회귀분석과 유사한 방법으로 반명제형 지식구조의 형태와 비슷하다. 둘째, 먼저 각 변수의 임계치(Threshold)를 찾아 결과값이 그 임계치에 해당하는지를 살펴 결정하는 일종의 의사결정트리(Decision Tree)를 찾아 결정하는 Non-compensatory 전략으로 명제형 지식구조와 흡사한 것을 알 수 있다. 이러한 두 가지 전략을 병행하여 사용하는 것을 혼합전략으로 칭하였는데 과거의 많은 연구들은 주어진 과제가 복잡하거나 시간의 제약이 있으면 인간은 Compensatory 전략에서 Non-compensatory 전략으로 옮겨가는 것을 보여줬다(Brehmer, 1987; Hagaford & Breimen, 1983; Kattan et al., 1993; Olshavsky, 1979; Paquette & Kida, 1988; Rothstein, 1986).

이러한 연구 중 특히 여러 개의 입력변수와 확률값으로 표현되는 결과가 있는 경우를 Multiple Cue Probability Learning(MCPL)로 명명하며 관련된 변수의 선정과 그 변수의 정보를 종합하여 의사결정을 내리는데 영향을 주는 것에 관한 많은 연구가 수행되었다. 예를 들면 과제의 난이도, 시간의 제약, 피드백의 유무, 추가정보의 사용, 전문가지식의 정도, 선형관계와 비선형관계, 판단의 일관성, 과거 자료의 사용 등이 최종의 사결정에 어떠한 영향을 주는가에 대한 연구가 수행되어왔다(Nutt, 1998). 본 연구는 이러한 과제의 난이도, 시간의 제약, 피드백의 유무 등은 연구범위에 포함하지 않았다. 그러나 추가정보의 사용은 설문에 전체 자료의 특성을 제시하여 고려하고, 전문가 지식의 정도는 전문가 선정에서 최대한 고려하고, 선형과 비선형의 과제는 명제형 지식구조와 비명제형 지식구조의 구도

로, 판단의 일관성은 AHP의 일관성으로 측정하여 본 연구에 포함하였다.

또한 Bolger & Wright(1994)는 기존의 전문가 판단에 대한 20여개의 실증적 연구를 정리하여 실제 전문가의 지식정도의 척도로서 생태적 유효성(Ecological Validity)과 의사결정에 대한 피드백의 유무에 따른 학습성(learnability)의 두 차원으로 나누어 분류한 후, 상충된 결과를 보이는 원인을 여러 가지 각도에서 분석하였다. 이러한 그의 연구결과 본 연구에서 사용하고자 하는 부도예측의 경우는 생태적 유효성은 높은 반면에 학습성은 낮아 부도에 대한 전문가의 판단은 낮아질 것으로 결론 지었다. 따라서 본 연구는 데이터 마이닝의 지식 표현에 대한 전문가의 신뢰도를 측정하고자 할 때, 전문가 판단의 변화는 측정하되 그 성과는 본 연구의 범위에 포함하지 않았다.

Dilla & Stone(1997)은 위험분류에 있어 의사결정지원의 형태로 숫자와 문자의 두 형태를 제시하고 위험에 대한 판단에서 문자로 표현되는 형태가 숫자로 표현되는 형태보다 변수의 가중치 부여와 최종판단에서 더욱 우수한 결과를 보인다는 연구결과를 발표하였다. 또한 우춘식 외 2인(1997)과 김광용 외 2인(1998)에서는 AHP를 이용하여 전문가의 일관성을 측정한 후, 전문가 지식의 판단 예측력을 평가하였는데, 일관성의 정도에 따라 각 변수의 가중치 및 예측력의 차이가 있는 것으로 나타났다. 따라서 본 연구에서는 여러 가지 데이터 마이닝 기법에 대한 전문가의 신뢰도를 측정함에 있어 자료의 특성(정량적 대 정성적)과 전문가 지식의 일관성(AHP 일관성)을 통제 변수로 사용하였다.

1 실제 본 연구에서도 전문가의 부도예측력은 데이터마이닝 모델 예측력보다 낮았다

3. 연구방법

본 연구의 연구방법은 실제 경영현장의 전문가를 대상으로 설문지와 인터뷰를 사용한 실험 연구(Experimental Design) 방법으로, 수집된 중소기업의 건전 및 부도위험자료를 무작위로 분석표본(Training set)과 검증표본(Test set)으로 분류해서, 여러 가지 데이터마이닝 기법들을 사용한 모형을 개발하고, 이러한 모형들을 전문가에게 제시하여 전문가의 신뢰도를 측정하였다.

본 연구에서는 실험실에서 주어진 모의화된 문제를 갖고 실험하기보다 실제 과거에 일어났던 부도와 건전경우를 가지고 경영현장의 전문가의 판단에 대한 실험을 하고자 한다. 이러한 실험 방법은 인간 판단의 유효성(performance)을 검증할 때, 인위적인 조작이 없어 생태적 유효성을 높이는 방법으로 간주되고 있다(Burn & Wright, 1991, Bolger & Wright, 1994).

또한 많은 모형을 동시에 보게되면 학습효과(Learning Effects)가 있어 실험결과에 편의를 주게 되므로 여러 전문가를 무작위로 나누어 각 전문가는 한가지 모형만을 이용하게 하였다. 즉 각 전문가는 모형의 도움없이 부도위험의 판단을 내리고, 또한 모형의 도움하에서의 부도위험을 판정하고, 마지막으로 왜 그런 결정을 내리는 지에 대한 설명을 본 후 최종 의사결정을 하게 하였다.

본 연구에서 사용한 데이터 마이닝기법²은 명제형 지식구조로는 CHAID(Chi-squared Automatic Interaction Detection)를 사용하였고, 반명제형 지식구조로는 다변량 판별분석(Multiple Discriminant Analysis : MDA), 비명제형 지식구조로 인공신경망(Artificial Neural Network : NN)을 이용하였

2 CHAID는 Answer Tree를, MDA는 SPSS를, NN은 Neuro Shell을 사용하였다.

으며, 자료특성별로 정량자료는 재무정보를, 정성정보는 비재무정보를 이용하였다. 따라서 지식구조(3그룹)*자료특성(2그룹)의 총 6그룹의 실험집단이 구성되었다.

3.1 데이터마이닝 모형의 개발 및 전문가 선정

3.1.1 데이터마이닝 모형개발

데이터 마이닝 기법별, 자료특성별 예측치를 추정하기 위하여 96-97년 사이의 중소기업을 대상으로 건전기업 10,600개와 부도기업 372개의 기업을 1차로 수집하였고, 자료정제작업을 통해 건전기업 7,636개와 부도기업 263개를 2차로 선정하였으며, 최종적으로 건전기업 262개와 부도기업 263개를 선정하였다.³ 즉, 최종표본 525개의 기업에 대해 모형 개발을 위한 분석표본은 부도기업 131개와 건전기업 131개로 구성하였으며, 개발된 모형의 검증을 위해 검증표본도 부도기업 132개와 건전기업 131개로 구성하였다.

(표 3) 분석에 사용된 정보

재무정보	
안정성	자기자본비율(%) 고정장기적합률(%) 당좌비율(%) 차입금의존도(%)
수익성	총자본경상이익율(%) 매출액영업이익율(%)
활동성	총자본회전율(회) 영업자산회전율(회)
생산성	설비투자효율(%) 부가가치율(%)
성장성	유형고정자산증가율(%) 매출액증가율(%)

3 representative sampling과 balanced sampling의 영향을 통제하기 위하여 본 연구에서는 건전기업과 부도기업이 선정될 비율을 동일하게 적용하는 balanced sampling을 하였다.

비재무정보	
사업성	성장전망(%) 수익전망(%) 판매전망 및 시장개척력 업종유망성(제품특성)
경쟁력	인력개발 기술개발 및 품질혁신 가격경쟁력 국제경쟁력
경영능력	경영자의 경영능력 노사관계 근로조건 및 복지수준
신뢰성	은행거래 신뢰성 세평

[표 3]은 분석에 사용된 정보를 나타내고 있으며, 재무정보는 안정성, 수익성, 활동성, 생산성, 성장성의 5개 요인에 12개 변수가, 비재무정보는 사업성, 경쟁력, 경영능력, 신뢰성, 기타의 5개 요인에 18개 변수가 선정되었다. 재무정보로 사용된 변수들은 비율 또는 연속변수이었으며, 비재무정보 중 성장전망, 수익전망은 연속변수이었고 나머지 변수는 5점 척도로 된 서열변수였다.

선정된 총 526개의 기업에 대해 MDA, CHAID, NN을 이용하여 신용평가 추정치(부도확률 및 부도가가능성)를 구하였다. CHAID 모형은 주어진 자료에 대한 최적의 IF-THEN 규칙 또는 의사결정트리를 형성함으로 분석표본을 통해 생성된 IF-THEN 규칙이 검증표본에서 부도인지 건전인지 판별 못하는 경우가 있다. 본 연구에서도 검증표본에 대해 판별불가의 표본이 재무정보일 경우 8개, 모든정보일 경우는 13개로 나타났으며, 비재무정보일 경우는 없는 것으로 나타났다. 본 연구에서 발생한 이러한 현상은 ID3를 이용해서 도산예측모형간의 예측력 비교를 한 많은 선행연구들 중에서 김광용 외 2인(1998)에서도 지적되고 있다. 따라서 CHAID, ID3 등의 귀납적 학습방법 모형을 이용해서 모형간 예측력 비

교를 한 선행연구는 한번쯤 고려해보아야 할 것으로 생각된다. 본 연구에서는 판별불가의 경우를 예측력에 포함한 경우(예측력 계산에서 제외)를 CH_1, 판별불가를 반영(잘못된 예측이라 간주) 한 것을 CH_2라 명명하여 별도로 정리하였다.

인공신경망은 역전파 알고리즘(Back-propagation Algorithm)을, 학습횟수는 3,000회~6,000회 사이로 하였고, 500회 이상 학습을 하였음에도 학습 성과가 없으면 학습을 그만두는 형태로 모형을 개발하였다. 은닉층(Hidden layer)의 노드(node)수는 여러 번의 시행을 거쳐 가장 높은 예측력을 높인 것을 이용하였는데, 재무정보에서는 7개, 비재무정보에서는 10개, 모든정보에서는 12개였다.

[표 4] 데이터 마이닝 기법의 예측 결과

		재무정보	비재무정보	모든정보
MDA	분석표본	69.5%	69.5%	72.9%
	검증표본	68.8%	65.8%	71.9%
CHAID	분석표본	75.0%	72.2%	80.0%
	검증 표본	CH-1 56.1% CH-2 54.4%	65.4%	61.6% 58.5%
NN	분석표본	72.1%	67.9%	70.6%
	검증표본	70.7%	69.6%	70.0%

[표 4]에서 보는 바와 같이 MDA를 이용한 신용평가 예측력은 재무정보와 비재무정보를 모두 이용한 모형(71.9%)이 가장 높은 예측력을 보였으며, 그 다음이 재무정보만을 이용한 모형(68.8%)이고, 비재무정보만을 이용한 모형(65.8%)이 가장 낮은 예측력을 보이고 있다. 분석표본과 검증표본에 따라 예측력에는 거의 차이가 없는 것으로 나타나고 있으나, 사용된 정보유형에 따라 큰 차이를 보이고 있다.

CHAID를 이용한 CH-1,CH-2의 예측력은 CH-1, CH-2 모두 비재무정보(65.4%)에서 가장 높은 예측력을 나타내고 있고, 분석표본과 검증표본의

예측력에 상당한 차이를 보이고 있다. 이러한 결과는 다른 귀납적 학습방법 모형과 마찬가지로, CHAID도 주어진 자료(분석표본)에 대해 최적의 의사결정트리를 형성하는 경향이 있으며, 따라서 안정성(robustness)이 떨어진다는 것을 알 수 있다. 그러나 CHAID와 같은 귀납적 학습방법 모형은 전문가에게 의존하는 바가 별로 없이 오직 수집된 자료만을 가지고 해당문제에 대한 지식을 획득 할 수 있기 때문에 수집된 자료의 질만 보장된다면 상당히 유용한 지식획득 방법이 될 수 있다는 점에서 지속적인 연구가 필요하다고 할 수 있다(이재식 & 한재홍, 1993). 인공지능경망의 예측력은 정보의 유형에 따라 예측력에 큰 변화가 없어 안정적(robustness)임을 알 수 있다.

3.1.2 전문가 선정

본 연구에서 선정한 전문가는 국내 신용평가기관의 신용평가 담당자 및 신용평가모형 개발자들로서 최소한 중소기업 신용심사 경력이 6년 이상 된 전문가를 선정하였다. 국내의 모든 신용평가기관에서 전문가를 고르게 선정하고자 하였는데, 그 내역은 신용보증기금에서 18명, 한국기업평가에서 6명, 한국신용평가에서 12명, 기술신용보증기금에서 6명, 한국신용정보에서 6명, 서울신용정보에서 12명을 합한 60명과 회계사 6명을 포함하여 총 66명의 전문가를 대상으로 실험을 하였다. 실험기간은 1999년 2월 1일부터 2월 13일 까지 12일 동안 직접방문을 통해 이루어졌다. 각 전문가는 무작위로 나뉜 설문(6가지 종류 중 1가지)을 받았으며, 설문을 완료한 전문가에게 도서상품권을 증정하여 설문참여와 설문결과에 질을 높이고자 하였다.

3.2 실험 설문지 구성 및 실험과정

3.2.1 실험에 사용될 표본기업 선정

여러 가지 데이터마이닝 기법을 통해 생성된 위험예측치에 대한 전문가들의 판단변화를 측정하는 것이 본 연구의 주된 목적이므로, 기본적으로 각 데이터마이닝 기법을 통해 생성된 위험예측치를 인식하기 전/후의 전문가 판단의 위험예측치 변화도를 살펴보게 된다. 따라서 각 데이터마이닝 기법별로의 예측력 차이에 의한 영향과 재무와 비재무 자료특성별 예측력 변화에 의한 영향을 최소화하기 위하여 가능한 데이터마이닝 기법별로 동일한 예측을 한 경우(MDA=CHAID=NN)와 자료특성별로 동일한 예측을 한 경우(정량적=정성적)를 표본선정의 기준으로 하였다. 또한 신용평가의 자료특성상 실제값(부도 vs 건전)과 예측값(부도 vs 건전)의 조합에 따른 영향력을 통제할 필요성이 있어 4가지 조합(실제부도→부도예측, 실제부도→건전예측, 실제건전→건전예측, 실제건전→부도예측)을 고려하여 실험에 사용할 표본의 기준점으로 선정하였다.

결론적으로 전문가를 대상으로 실험할 표본기업은 부도위험 예측력이 ① 최대한 MDA=CHAID=NN이고 ② 최대한 정량정보(재무정보)=정성정보(비재무정보)의 조건을 만족하는 기업을 1차로 선정한 후, 최종적으로 실제부도→부도예측, 실제부도→건전예측, 실제건전→건전예측, 실제건전→부도예측한 4가지 경우에 대해 각각 2개씩 무작위로 추출하여 총 8개의 표본기업을 선정하였다.

3.2.2 실험 설문지 구성

설문지 구성은 설문에 관한 인사말, 설문자료

에 대한 소개, 설문방법에 대한 상세한 안내를 첫 장에 하였다. 그 내용은 설문에 사용된 설문 기업의 산업군, 근거년도, 전체 기업의 표본수, 전체표본의 구성비율(건전:부도)⁴이고, 특히 틀리게 예측한 결과가 포함되어 있음을 강조하여 설문의 응답에 편이가 생기지 않도록 각 전문가의 소신있는 평가를 부탁하였다. 두 번째 장에 모집단의 특성을 제시하여 설문에 참고하도록 하였는데, 각 자료(정량, 정성)에 대한 소개, 의미, 적도와 신용평가예측에 사용된 526개 기업(건전 262개, 부도 263개)에 대한 평균값, 표준편차, 최소값, 최대값 등을 제시하였다.

세 번째 장에는 3개의 영역으로 나누어 편집하였는데 있는데, 한 영역의 답을 한 후 그 가려진 부분을 뜯어내고 그 다음 영역에 답을 하게 하는 방법을 사용하였다. 이러한 3 가지 영역의 첫째부분은 8개 기업에 대한 재무나 또는 비재무 변수별로 비율값 또는 연속값이 제시되어 있고, 그러한 자료만을 사용하여 전문가가 부도확률을 예측하게 하였다. 둘째 영역은 데이터마이닝 기법(MDA, CHAID, NN)중 한가지 방법을 이용하여 생성된 부도위험 예측치를 8개 기업에 대해 제시하였으며, 그 예측치를 보고 다시 한번 각 8개 기업에 대한 부도위험 확률을 수정하게 하였다. 마지막 부분은 각 데이터마이닝 기법이 제시하는 부도위험 예측치에 대한 이론적 설명(Rationale)이 제시되어 있으며, 이러한 설명을 보고 전문가는 다시 한번 부도위험 예측치를 추정하게 하였다.

마지막 장에는 본 연구에서 선정한 전문가들이 어느 정도 지식의 일관성을 가지고 있으며 또한 어떠한 자료에 더 많은 중요도를 부여하고 있는지를 조사하고자 전문가들이 사용한 재무, 비재무항목에 대한 AHP(Analytic Hierarchy Process) 조사를 실시하였다(김광용, 1998)

3.2.3 실험과정

상기 기술한 바와 같이 설문형태는 6가지가 있으며, 각 전문가는 6가지의 설문지 중 하나를 무작위로 배정 받았으며, 일정 시간의 사용이나, 기존 자료의 탐색, 참고자료의 사용 등에 관한 제약은 전혀 주지 않았다. 즉 완전히 자유로운 상황에서 전문가의 지식을 최대한 사용할 수 있도록 배려하였다. 대략적으로 설문을 완성하는데 평균 20-30분 정도가 소요되었으며 설문보조자는 옆에서 설문자의 질문에 즉시 응답할 수 있도록 충분한 교육을 받고 각 설문자의 질문에 만 답을 하도록 하였으며 실험과정은 아래와 같은 단계로 실시하였다.

- 1단계: 세 번째 장의 첫 번째 부분을 열면 8개의 기업에 대한 재무 또는 비재무정보가 제시되는데, 전문가들은 이를 분석한 후 각 기업의 부도위험 예측치를 결정한다.
- 2단계: 다음 부분에서는 동일한 8개 기업에 대한 각 데이터 마이닝 기법(MDA, CHAID, NN 중 하나)을 통해 생성된 부도위험 예측치가 제시되어 있는데, 이를 참고로 다시 각 기업에 대한 부도위험 예측치를 결정한다.
- 3단계: 마지막 부분에서는 각 데이터 마이닝 기법이 제시한 부도위험 예측치에 대

4 예비실험결과 전문가들은 실제 모집단의 부도와 건전의 발생비율(사건확율)을 고려하여 부도위험 예측치를 평가하고 있는 것으로 밝혀져, 본 연구에서는 balanced sampling(부도:건전=50:50)임을 강조하여 사전확율을 고려한 판단의 편의를 제거하고자 하였다.

한 이론적 근거를 설명하고 있다. 예를 들면 CHAID+비재무인 경우 비재무 정보에 대한 IF-THEN 규칙을 제시하고 있으며, MDA+재무인 경우 재무변수에 대한 계수값과 합산과정을, NN의 경우는 NN에 대한 간략한 소개를 하고 있다. 이를 참고로 하여 전문가들은 다시 8개 기업에 대한 부도위험 예측치를 결정하게 된다.

살펴보면 부도 예측력과 안정성에서 기하평균이 산술평균보다 우수한 바를 입증하고 있고, 일관성이 높은 전문가가 더욱 높은 부도예측력을 보이고 있으므로, 기하평균의 자료에 기인하여 볼 때 일관성이 높은 전문가(CR<0.1)는 부도예측에서 정성적 정보에 더욱 많은 가치를 두고 있고 부도예측력도 높을 것이라는 것을 시사한다고 하겠다.

4. 실험결과

4.1 전문가들의 자료에 대한 상대적 중요도

[표 5]는 AHP를 이용한 신용평가 전문가들이 생각하는 자료의 특성에 따른 상대적 가중치를 나타내고 있다. 일관성 비율(CR: Consistency Ratio)이 0.1 이하인 일관성이 높은 전문가는 전체의 약 23%로 나타났으며 특히 0.2이하인 전문가가 전체의 80%를 차지하여 전반적으로 타 연구에 비하여 본 연구의 전문가는 높은 일관성을 갖고 있는 것으로 나타났다. 일관성비율(0.1)의 구분에 따라 정량적 정보의 가중치와 정성적 정보의 가중치는 산술평균에서는 거의 차이가 없음을 볼 수 있다. 또한 이러한 가중치의 종합을 보면 일관성 비율에 관계없이 정량적 정보보다는 정성적 정보에 더 많은 가중치를 부여하고 있는 것으로 나타나고 있다.

그러나 기하평균의 경우에는 전문가의 일관성 비율별로 많은 차이가 나서 향후 다수의 전문가 지식을 종합하는데 보다 깊은 연구가 필요한 것으로 보인다. 단 기존의 연구 중 우춘식의 2인(1997)의 연구와 김광용 외 2인(1998)의 결과를

[표 5] 자료특성별에 따른 전문가 가중치

위험예측 변수	산술평균		기하평균		
	전체	CR<0.1	전체	CR<0.1	
정량적 재무정보	안정성	14.1	15.4	14.5	13.5
	수익성	13.2	14.2	15.2	11.3
	활동성	5.5	5.7	5.6	4.1
	생산성	6.9	4.7	7.1	3.5
	성장성	6.9	6.4	6.9	4.8
	합계	46.6	46.4	49.3	37.2
정성적 비재무정보	사업성	11.4	11.8	11.2	13.1
	경쟁력	13.3	13.1	12.7	13.5
	경영능력	12.8	12.9	11.4	14.9
	신뢰성	11.2	10.9	10.6	14.0
	기타	4.8	5.3	4.8	7.2
	합계	53.5	54.0	50.7	62.7

4.2 데이터마이닝 기법별, 실험 단계별 전문가들의 부도위험 예측치

[표 6]은 데이터마이닝 기법이 생성한 각 기업의 부도위험 예측치와 각 단계별 전문가들이 예측한 각 기업의 예측치를 나타내고 있다. 대체적으로 전문가들은 각 기업의 정보만을 고려하였을 경우(1단계) 각 기업에 대한 부도위험 예측치를 낮게 평가하는 경향이 있는 것으로 나타나고 있으나, 데이터 마이닝 기법이 제시하는 각 기업의 부도위험 예측치를 인식한 후에는(2

단계) 부도위험 예측치가 높아지는 것을 볼 수 있다. 특히 데이터마이닝 모델이 제시하는 이론적 근거를 보고 난 후에는(3단계) 그 예측치가 조금 더 증가하는 것을 볼 수 있다. 또한 정량 정보(재무정보)에 대한 변화보다는 정성정보(비재무 정보)에 대해 더 큰 변화를 보이고 있어 [표 5]에서 나타난 정성정보에 대한 가중치를 높게 두는 것과 관련성이 있음을 알 수 있다.

[표 6] 데이터 마이닝 기법별, 실험 단계별 위험 예측치

자료 특성별	데이터 마이닝 기법별	1단계 (exp1)	2단계 (exp2)	3단계 (exp3)	Computer's Estimate
정량적 재무정보	MDA	32.14	38.53	38.73	49%
	CHAID	34.11	35.76	36.81	47%
	NN	37.95	39.00	39.76	43%
정성적 비재무 정보	MDA	44.02	46.58	47.75	47%
	CHAID	32.56	38.67	47.55	42%
	NN	43.72	43.55	43.26	45%

그러나 이러한 해석은 상당히 조심스럽게 접근하여야 하는데 그 이유는 실제 주어진 표본기업 중에는 부도와 건전이 혼합되어 있어 건전의 경우에 전문가의 지식이 반영되면 그 예측치가 낮아지고 부도의 경우에는 그 예측치가 높아져 전체의 평균예측치가 상쇄되는 효과가 나타나기 때문이다. 따라서 각 데이터마이닝 기법이 제시하는 지식에 대한 전문가의 신뢰도는 다음장에서 제시될 전문가 판단의 변화량으로 측정하는 것이 타당한 방법일 것이다.

4.2 실험단계별 부도위험 예측치의 변화량

[표 7]은 실험 단계별 전문가들의 부도 위험 예측치의 변화량을 나타내고 있다. 즉, 각 기업에 대한 데이터 마이닝 기법의 부도위험 예측치

를 인식하기 전(exp1)과 인식한 후(exp2)의 부도 위험 예측치의 순수 변화량(|Differ1|=exp1-exp2), 데이터 마이닝 기법의 부도위험 예측치에 대한 논리적 설명을 인식한 후(exp3)의 부도위험 예측치의 순수변화량(|Differ2|=exp2-exp3), 데이터 마이닝 기법의 부도위험 예측치와 논리적 설명을 인식한 후의 부도위험 예측치의 순수 변화량(|Differ3|=exp1-exp3)를 나타내고 있다.

[표 7] 실험단계별 부도위험 예측치의 변화량

자료 특성별	데이터 마이닝 기법별	Differ1 (exp1~exp2)	Differ2 (exp2~exp3)	Differ3 (exp1~exp3)
정량 정보	MDA	8.69 (13.90)	4.58 (6.87)	11.36 (13.84)
	CHAID	4.60 (7.72)	8.00 (12.37)	8.90 (12.48)
	NN	5.89 (7.23)	5.47 (9.03)	9.38 (12.35)
정성 정보	MDA	9.13 (16.22)	3.65 (6.03)	9.75 (13.90)
	CHAID	7.82 (11.81)	12.83 (16.34)	18.60 (18.15)
	NN	4.71 (8.75)	3.27 (6.99)	7.18 (9.92)

* 괄호안은 표준편차임

|Differ1|의 경우를 보면 자료특성별로는 큰 차이를 보이고 있지 않으며, 데이터 마이닝 기법별로 보면 MDA에서 많은 변화가 있음을 볼 수 있다. 그러나 |Differ1|의 변화량은 데이터 마이닝 기법에 의한 변화이기보다는 각 데이터 마이닝 기법이 제시하는 부도위험 예측치의 차이에 의한 결과라 할 수 있다. 즉, [표 6]에서처럼 MDA에 의한 각 기업의 부도위험 예측치가 다른 데이터 마이닝 기법을 통한 부도위험 예측치 보다 높게 제시되어, 이에 따라 CHAID, NN보다 더 많은 변화를 보인 결과라 할 수 있다. 또한 |Differ1|의 경우는 전문가가 어떤 기법을 사용하여 부도 예측 위험치가 도출된 것인지를 전혀 모르고 단지 데이터마이닝 기법중의 하나를 사용한 컴퓨터 모델링의 결과만 알고 있는 상황에서의 판단

의 변화이므로 데이터마이닝 기법의 영향에 의한 것으로 볼 수 없고 단지 컴퓨터 결과에 대한 전문가의 판단 변화로 볼 수 있다.

|Differ3|의 경우는 데이터 마이닝의 도움없이 판단한 후, 컴퓨터 결과만 보고 그 결과가 도출된 논리적 이유를 본 후의 전문가 판단의 변화이다. 이러한 변화 역시 순수한 데이터 마이닝 기법으로부터 도출된 지식에 의한 영향이라고 보기가 힘든데 그 이유는 |Differ3|에는 |Differ1|의 영향이 그대로 남아 비록 사용된 데이터마이닝 기법의 논리적 설명을 본 후에 전문가의 판단을 변화했다고는 하지만 앵커링조정법(Anchoring and adjustment) 영향이 그대로 남아있을 소지가 높기 때문이다. 또한 중간단계(2단계)의 수정 후에 최종단계(3단계)의 수정이 이루어져 절대변화량의 측정에도 문제가 있어 데이터마이닝 기법의 사용에 의한 전문가의 판단 변화를 측정하는 데는 무리가 있다.

|Differ2|의 경우는 먼저 데이터마이닝의 부도 예측치를 보고 결정한 판단과 각 데이터 마이닝의 이론적 설명을 보고 난 후의 판단간의 변화량으로 실제 데이터마이닝 기법의 사용에 의한 영향으로 간주할 수가 있으며, 특히 단계1에서 이미 컴퓨터 결과에 의한 영향력을 고려하였으므로 |Differ2|는 각 데이터마이닝 기법의 지식에 기인한 순수한 전문가 판단의 변화량으로 간주할 수 있으며, 따라서 데이터마이닝의 지식에 대한 전문가의 신뢰도 측정에 가장 이상적인 측정치라 할 수 있다. 본 연구에서는 |Differ2|를 종속변수로 선정하여 전문가 신뢰도의 측정치로 사용하였다.

4.3 데이터마이닝 기법에 따른 전문가 판단의 변화

[표 8] |Differ2|에서의 전문가 판단의 변화

DSS Type	정보유형	AHP Group	Mean	Standard Deviation	Sample Size
MDA	정량적 재무정보	CR<=0.1	0.00	0.00	16
		CR>0.1	5.70	7.46	64
		Total	4.56	7.05	80
	정성적 비재무정보	CR<=0.1	0.87	2.23	40
		CR>0.1	5.96	7.13	48
		Total	3.65	6.02	88
	Total	CR<=0.1	0.62	1.92	56
		CR>0.1	5.81	7.29	112
		Total	4.08	6.52	168
CHAID	정량적 재무정보	CR<=0.1	12.70	16.42	24
		CR>0.1	6.23	10.06	64
		Total	8.00	12.36	88
	정성적 비재무정보	CR<=0.1	21.56	16.70	16
		CR>0.1	10.89	15.72	72
		Total	12.82	16.34	88
	Total	CR<=0.1	16.25	16.89	40
		CR>0.1	8.69	13.51	136
		Total	10.41	14.65	176
NN	정량적 재무정보	CR<=0.1	5.87	9.73	16
		CR>0.1	5.37	8.93	72
		Total	5.46	9.02	88
	정성적 비재무정보	CR<=0.1	1.25	2.31	8
		CR>0.1	3.46	7.26	80
		Total	3.26	6.98	88
	Total	CR<=0.1	4.33	8.27	24
		CR>0.1	4.36	8.12	152
		Total	4.36	8.12	176

위의 [표 8]을 살펴보면 MDA의 경우 재무나 비재무 정보에 관계없이 일관성이 높은 전문가(CR<=0.1)는 MDA의 부도예측치에 대한 설명을 보고도 큰 변화가 없는 것을 알 수 있다. 또한 NN의 경우에 전문가의 일관성과는 관계없이 재무나 비재무 정보에서 비슷한 변화를 보이고 있는 것을 알 수 있다. 그러나 CHAID의 경우, 일관성이 높은 전문가는 CHAID가 제시하는 규칙(Rule)을 보고 보다 많은 신뢰를 하여 변화의 폭이 큰 것을 알 수 있으며, 특히 비재무정보의 경우는 그 변화의 폭이 재무정보의 경우보다 약 2배 가까이 큰 것으로 나타났다. 즉 전문가는

여러 가지 데이터 마이닝이 제시하는 지식에 대하여 갖는 신뢰도가 다르다는 것을 알 수 있으며, 특히 이러한 차이는 전문가 지식의 정도(일관성)와 제시되는 자료의 특성별로 차이가 있다는 것을 알 수 있다. 따라서 이러한 데이터 마이닝 기법별로 어떠한 차이가 있는지 또 데이터 마이닝 기법과 자료특성, 전문가의 일관성과는 어떠한 상호작용이 있는지를 알아보기 위하여 Full Factorial Design을 사용한 GLM(Generalized Linear Model) 방법⁵을 이용하여 분석하였다.

[표 9] 데이터마이닝 기법에 따른 전문가 판단의 변화

자료	Type III Sum of Square	Degree of Freedom	Mean Square	F_Ratio	Sig.
Corrected Model	8982.21	11	816.56	8.02	0.000
Intercept	13536.64	1	13536.64	132.95	0.000***
DSS	6698.67	2	3349.34	32.89	0.000
DATA	139.19	1	139.19	1.36	0.243
AHP-Group	45.82	1	45.82	0.45	0.503
DSS*DATA	1261.85	2	630.92	6.19	0.002***
DSS*AHP-Group	3097.21	2	1548.60	15.21	0.000***
DATA*AHP-Group	9.42	1	9.41	0.09	0.761
DSS*DATA*AHP-Group	141.58	2	70.79	0.69	0.499
Error	51721.15	508	101.81		
Total	81481.00	520			
Corrected Total	60703.37	519			

*** : P<.01

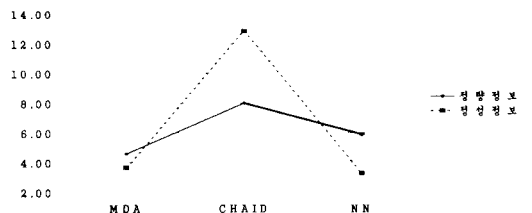
[표 9]에서 보는 바와 같이 데이터마이닝 기법(DSS)에 따라 전문가 판단의 변화는 통계적으로 유의한 차이가 있는 것으로 나타났으나, 자료의 특성(DATA) 및 전문가 지식의 일관성(AHP-Group)과는 유의한 차이가 없는 것으로 밝혀졌다. 즉 여러 가지 데이터마이닝 기법의 지식에 따라 전문가의 신뢰도는 차이가 있는 것으로 밝혀졌다. 따라서 어떤 데이터마이닝 기법의 신뢰도가 다른지를 확인하기 위하여 Tukey의

LSD를 이용하여 사후다중비교(Multiple Comparison Test)를 하였다. 분석한 결과를 정리한 [표 10]에서 보면 MDA와 NN의 집단은 CHAID와 통계적으로 유의한 차이를 보이고 있다.

[표 10] 데이터마이닝 기법의 사후다중비교

데이터마이닝기법	N	Subset	
		1	2
MDA	168	4.08	
NN	176	4.36	
CHAID	176		10.41

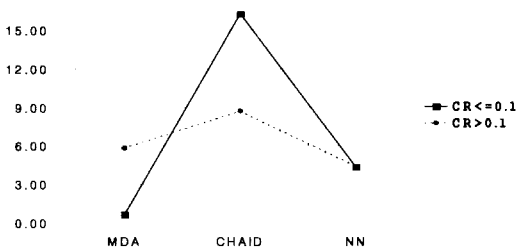
따라서 전문가들은 비(반)명제형 지식구조를 갖는 MDA와 NN보다는 CHAID와 같은 명제형 지식구조에 더욱 더 많은 신뢰를 주어서 전문가의 부도예측치의 판단에서 더욱 큰 변화를 보인다고 할 수 있다. 이러한 실증적 결과는 매우 중요한 것으로써 이미 많은 기존의 연구가 강조했던 귀납적 학습방법의 설명력은 단지 이해를 돕는 정도가 아니라 전문가의 신뢰도에도 깊은 영향을 미친다는 것을 알 수 있다. 따라서 데이터마이닝 기법을 활용하여 전문가의 의사결정을 돕고자 할 때는 이런 명제형 지식구조를 고려한 유저인터페이스가 설계되어야 하며, 특히 각 데이터마이닝 모형의 장점을 고려한 혼성모형(Hybrid Model)의 개발에는 단순히 예측력의 성과뿐만 아니라 전문가의 신뢰도를 고려하여 혼성모형의 최종 지식표현은 명제형지식구조로 설계할 필요성을 알 수 있다.



[그림 1] 데이터마이닝과 자료특성간의 상호작용효과

5 SPSS를 사용하였다

또한 위의 [표 9]에서 보듯이 여러 가지 데이터마이닝 기법과 자료의 특성(DSS*DATA)은 상호작용이 있는 것으로 나타났다. 이러한 상호작용효과가 어떻게 발생되었는지를 정리한 [그림 1]을 살펴보면, 명제형 지식구조 CHAID는 자료의 특성에 관계없이 전문가의 신뢰를 많이 받아 가장 많은 판단의 변화가 이루어지지만 특히 정성적 정보에서 더욱 많은 변화가 이루어지는 것을 알 수 있다. 그러나 반(비) 명제형 지식구조를 갖고 있는 MDA나 NN은 도리어 정량적 정보에서 더 많은 전문가 판단의 변화가 발생하여 이러한 상호작용이 발생하였던 것이다. 즉 NN이나 MDA의 경우 전문가는 정량적 정보에 더 많은 신뢰도를, CHAID의 경우에는 정성적 정보에 더 많은 신뢰도를 갖고 있는 것을 알 수 있다. 본 연구결과는 숫자로 표시한 의사결정지원도구가 전문가 판단의 일관성을 높인다는 Dilla & Stone(1997)의 연구와도 일맥상통하여 향후 보다 세밀하고 다양한 연구가 필요하다 하겠다. 특히 이러한 실증적 결과는 향후 데이터마이닝 기법을 활용하여 전문가의 지식을 돕고자 할 때 정보의 특성을 고려한 모델링이 필요하다는 것을 알 수 있다.



[그림 2] 데이터마이닝과 판단일관성간의 상호작용 효과

마찬가지로 데이터마이닝 기법과 전문가 지식의 일관성 유무(DSS*AHP-Group)사이에도 상호

작용이 있는 것으로 나타나고 있다. 그 이유는 [그림 2]에서 보듯이 CHAID의 경우가 가장 높은 신뢰도를 받아 전문가 판단의 변화가 많이 이루어지고 있으며 특히 일관성이 높은 전문가는 CHAID에 더욱 많은 신뢰를 하고 있다는 것을 알 수 있다. 그러나 MDA의 반명제형 지식의 경우에는 반대로 일관성이 높은 사람은 거의 신뢰를 하고 있지 않으며(변화량이 거의없음), 일관성이 낮은 전문가가 도리어 높은 신뢰도를 보여주고 있다. 흥미로운 것은 NN의 경우는 일관성의 차이가 전문가의 판단에는 거의 영향을 미치지 않아 NN의 설명력 부재가 전문가 지식의 일관성에 관계없이 동일하게 영향을 미치는 것을 알 수 있다.

5. 결론 및 향후 연구방향

본 연구는 여러 가지 데이터마이닝 기법의 지식에 대한 전문가의 신뢰도를 측정하고자 하였다. 특히 이러한 전문가의 신뢰도를 사용된 자료의 특성별, 전문가 지식의 일관성별로 조사하였다. 분석결과 명제형 지식구조를 갖고있는 CHAID는 전문가가 가장 신뢰하는 데이터마이닝 기법으로 밝혀졌으며 특히 정성적 정보를 사용하는 경우와 전문가의 지적 일관성이 높을수록 더욱 높은 신뢰를 보이는 것으로 나타났다. 즉, 인간사고의 논리구조와 흡사한 명제형 지식구조를 가지면서 일상언어와 같은 정성정보의 형태를 제시하는 것이 인간의 판단에 가장 큰 영향을 준다고 할 수 있다.

전문가 지식의 일관성 유무와 자료의 특성별로 전문가의 신뢰도가 유의한 차이를 보이지는 않았지만, 전문가의 판단은 CHAID의 경우에는 정성적 정보에, MDA나 NN은 정량적 정보에 더

많은 신뢰를 주는 것으로 밝혀졌다. 또한 CHAID의 경우에는 일관성이 높은 전문가가 MDA의 경우에는 일관성이 낮은 전문가가 더욱 많은 신뢰를 보이는 것으로 나타났으며, NN의 경우에는 차이가 없는 것으로 밝혀졌다. 이러한 연구 결과는 Dilla & Stone(1997)이 발표했던 문자로 표현되는 의사결정지원도구가 사용된 변수의 가중치 부여와 최종판단에 긍정적인 영향을 미친다는 연구결과와 흡사하며 본 연구는 그들의 연구를 좀더 체계적이고 구체적으로 밝혀주고 있다. 즉 현재의 여러 가지 데이터마이닝 기법을 전문가의 판단을 돕는 지능형의사결정지원시스템(IDSS: Intelligent Decision Support System)으로 발전시키고자 할 때는 이러한 데이터마이닝 기법의 특성별로, 사용되는 정보의 특성별로, 전문가와 비전문가의 특성별로 차이를 고려한 유저 인터페이스의 설계가 필요하다는 것을 구체적으로 보여주고 있다(Lamberti & Wallace, 1990, Lim & O'Connor, 1996).

또한 여러 가지 데이터 마이닝 기법의 장점을 혼합한 혼성모형(Hybrid Model)의 개발에도 이러한 데이터마이닝 기법에 대한 전문가의 신뢰도를 고려하여 설계할 필요가 있다는 것이다(김광용, 1998). 현재 많은 데이터마이닝 연구가 예측력의 향상을 주로 고려하고 있는데 모델의 성과는 예측력, 설명력, 안정성, 단순성 등 사실상 고려해야 하는 많은 측정치가 있으며 본 연구에서도 제시하듯이 모델의 설명력은 전문가의 신뢰도에 유의한 영향을 미치므로 가능한 혼성모형의 최종지식구조는 명제형 지식구조인 IF-THEN 구조가 바람직하다는 것을 알 수 있다.

향후 연구의 과제는 앞서서도 지적하였듯이 피드백의 영향, 과제의 난이도 영향, 의사결정 시간의 제약 등 기존 의사결정이론에서 주로 많

이 연구해은 여러 통제변수의 영향을 고려한 좀더 세밀한 연구가 필요하다 하겠다. 특히 이러한 연구는 실제 급변하는 경영현장에서 의사결정을 수시로 내려야 하는 전문가의 판단을 돕는데 매우 중요한 연구가 될 것이다. 마지막으로 본 연구를 좀더 확장하여 데이터마이닝 기법의 지식에 관한 전문가의 신뢰도를 측정 한 후 이러한 신뢰도가 실제 예측력의 성과에는 어떠한 영향을 미치는지에 대한 연구가 필요하다 하겠다. 아쉽게도 본 연구는 예측력의 성과를 측정하기에는 부적절한 연구설계를 하고있어 이러한 해석이 어려웠는데, 예를 들면 설문자체가 부도에 대한 예측결과를 묻고있어 전반적으로 부도에 대하여 좀더 많은 가중치가 주어지는 프레이밍(Framing) 효과가 발생하는 문제가 있었다(Morwitz & Schmittlein, 1998). 또한 Bolger & Wright(1996)는 전문가 판단의 성과측정연구에는 판단에 대한 피드백이 매우 중요한 요소라고 지적하였는데 본 연구는 피드백을 연구설계에 포함하지 못하였다. 향후 데이터마이닝 기법을 활용한 전문가의 판단성과측정에 관한 보다 엄격히 설계된 많은 연구가 필요할 것으로 보인다.

참 고 문 헌

- 김광용, "여러가지 Inductive 방법에 대한 통합모델 개발과 그 실증적 유효성에 대한 연구", *한국경영과학회지*, 제23권 3호(1998, 9), pp.185-207
- 김광용·손광기·홍은선, "Data Mining 기법들과 전문가들로부터 추출된 지식에 관한 실증적 비교연구", *한국전문가시스템학회지*, 제4권 1호(1998, 6), pp.41-58
- 우춘식·김광용·강성범, "LOGIT 분석과 AHP 분석을 이용한 부도예측모형의 비교연구",

- 재무관리연구, 제14권 2호(1997), pp.229-252
- 이재식, 한재홍 “NN을 이용한 중소기업 도산예 측에 있어서의 비재무 정보의 유용성 검증”, *한국전문가시스템학회지*, 제1권 1호(1995), pp.123-134
- Arizne, B., and P. N. Narasimha, “An experimental investigation of predictive accuracy of induction and regression”, *Expert Systems with Applications*, Vol. 7 (1994), 535-544.
- Bolpus, F and G. Wright, “Assessing the quality of expert judgement: Issues and analys”, *Decision Science*, Vol. 11(1994), 1-24.
- Brehmer, B. “Note on subject’s hypotheses in multiple-cue probability learning”, *Organizational Behavior and Human decision Process*, Vol. 40(1987), 323-329
- Bunn, D. and G. Wright, “Interaction of judgemental and statistical freacasting methods: Issues and Analysis”, *Management Sciesnce*, Vol. 37(1991), 510-518
- Casey, C. and Selling, T. I “The Effect of Task Predictability and Prior Probability Disclosure on Judgment Quality and Confidence”, *Accounting Review*, Vol. 61(1986), 302-317.
- Chalos, P. “The Superior Performance of Loan Review Committee”, *Journal of Commercial Bank Lending*, Vol. 68(1985), 60-66.
- Chung, H. M., and Silver, M. S, “Rule-based expert systems and linear models: An empirical comparison of learning-by-examples methods”, *Decision Sciences*, Vol. 23(1992), 687-707.
- Bell, D. E., H. Raiffa and A. Tversky, “Decision Making : Descriptive, Normative and prescriptive interactions”, *Cambridge University Press*, Cambridge, 1988.
- Dawes,R.M. and Corrigan, B. “Linear Models”, *Psychological Bulletin*, Vol. 81(1974), 95-106
- Dilla, W. N. “Representations as decision aids: The asymmetric effects of words and numbers on auditor’s inherent risk judgment”, *Decision Science*, Vol. 28, No. 3(1997), 709-743.
- Gim, G., “Hybrid Systems for Robustness and Perspicuity: Symbolic Rule Induction Combined with a Neural Net or a Statistical Model”, *Doctoral Dissertation*, GSU,(1995)
- Gim, G., and Whalen, T. “Dimensions of knowledge: Facts or skills, words or numbers”, *The Proceedings of North American Fuzzy Information Proceeding Society*, (1994),447-448.
- Hagaford, R., and Breimen, B. “Does Having to justify one’s judgements change the nature of the judgement process?”, *Organizational Behavior and Human Performance*, Vol. 31(1983), 323-329
- Hart, A.. “Knowledge Acquisition for expert systems (2nd Ed.)”, *New York:Macgraw-Hill*, (1992)
- Huss, W. R. “Comparative Analysis of Company Forecasts and Advanced Time-Series Techniques Using Annual Electric Utility Energy Sales Data”, *Journal of Forecasting*, Vol. 1(1985), 217-239.
- Kattan, M. W., Adams, D. A., and Parks, M. S. “A comparison of machine learning with human judgment”, *Journal of management Information Systems*, Vol. 9, No. 4(1993), 37-57.
- Kim, C. N., H. M. Chung and D. B. Paradise., “Inductive modeling of expert decision making in loan evaluation: a decision strategy perspective”, *Decision Support System*, Vol. 21(1997), 83-98.
- Langley, P and H. A. Simon, “Application of machine

- learning and rule induction”, *Communication of the acm*, Vol. 38, No. 11(1995), 55-64.
- Lee, H. Y, “Combining Judgements for Better Decision : A Study for Investigating Effective Combining Schemes”, *Journal of the Korean OR/MS Society*, Vol. 21, No. 3, (1996), 159-174.
- Liang, T. P. “A composite approach to inducing knowledge for ES design”, *Management Science*, Vol. 38(1992), 1-17.
- Liang, T. P., Chandler, J. S., and Han, I. “Integrating statistical and inductive learning methods for knowledge acquisition”, *Expert Systems with Applications*, Vol. 1(1990), 391-401.
- Liang, T. P., Chandler, J. H., Han, I., and Roan, J. “An empirical investigation of some data effects on the classification accuracy of probit, ID3, and neural networks”, *Contemporary Accounting Research*, 9(1992), 306-328.
- Libby, R., “Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence”, *Journal of Accounting Research*, Vol. 23(1975), 150-161.
- Lim, J. S and m. O’Connor, “Judgmental forecasting with interactive forecasting support systems”, *Decision Support Systems*, Vol. 16(1996), 339-357.
- Mary, Z, “Decision Making”, *SAGE*, California, 1992.
- Mathews, B. P. and Diamantopoulos, A, “Judgmental Revision of Sales Forecasts: A Longitudinal Extension”, *Journal of Forecasting*, Vol. 8(1989), 129-140.
- Morwitz, V. G. and Schmittein, D. C, “Testing new direct marketing offerings: The interplay of management judgment and statistical models”, *Management Science*, Vol. 44, No. 5(1998), 610-628.
- Newell, A. and Simon, H. “Human Problem Solving”, *Englewood Cliffs, NJ:Prentice Hall*, (1972)
- Nutt, P. C, “How decision makers evaluate alternatives and the influence of complexity”, *Management Science*, Vol. 44, No. 8(1998), 1148-1166.
- Olshavsky, R.W. “Task complexity and contingent processing in decision making: A replication and extension”, *Organizational Behavior and Human Performance*, Vol. 24(1979), 300-316
- Paquette, L. and Kida, T. “The effect of decision strategy and task complexity on decision performance”, *Organizational Behavior and Human Decision Process*, Vol. 41(1988), 128-142
- Payne, J.W. “Task complexity and contingent processing in decision making: An Information Search and protocol analysis”, *Organizational Behavior and Human Performance*, Vol. 16(1976), 366-387
- Quinlan, R. “Induction of decision trees”, *Machine Learning*, Vol. 1(1986), 81-98.
- Reinmuth, J. E. and Guerts, M. D. “A Bayesian Approach to Forecasting Efforts of Atypical Situations”, *Journal of Marketing Research*, (1972).
- Rothstein, H. G. “The effect of time pressure on judgement in multiple que probability learning”, *Organizational Behavior and Human Decision Process*, Vol. 37(1986), 83-92
- Shaw, M. J. “Machine learning methods for intelligent decision support”, *Decision Support Systems*, Vol. 10(1993), 79-83
- Shepanski, A. “Test of Information Processing Behaviour in Credit Judgement”, *Accounting Review*, Vol. 58(1983) , 581-599
- Turban, E and J. Aronson. “Decision support systems and intelligent systems”, 5th, *Prentice-Hall*, New

- Jersey, 1998.
- Whalen, T. and Gim, G. "Second Order Logical System for Risk Classification In a Newly Developed Country", *International Journal of Uncertainty, Fuzziness, and Knowledge Based Systems*, Vol. 4(1996), 421-430.
- Whitred, G. and Zimmer, I. "The Implication of Distress Prediction Models for Corporate Lending", *Accounting and Finance*, Vol. 25(1985) , 1-13.
- Wolfe, C. and Flores, C. "Judgemental Adjustment of Earnings Forecasts", *Journal of Forecasting*, Vol. 9 (1990), 389-406.
- Woong B.K., Bodnovich T.A., & Selvi Y., "Neural Network Applications in Business: A review and Analysis of the literature(1988-1995)," *Decision Support System*, Vol.19, No.4(1997), pp.301-320
- Woong B.K. & Selvi Y., "Neural Network Applications in Finance: A review and Analysis of the literature(1990-1996)," *Information & Management*. Vol. 34(1988), 129-139
- Zeleny, M. "On the inadequacy of the regression paradigm used in the study of human judgment", *Theory and Decision*, Vol. 7(1976), 57-65
- Zey, M. "Decision Making: Alternatives Rational Choice Model," *Sage publications*, Newbury, 1992
- Zopounidis C., Doumpos M., & Matsatsinis N.F., "On the Use of Knowledge Based DSS in Financial Management:A Survey," *Decision Support Systems*, Vol.20, No.2(1997), pp.259-277