

리프집합을 이용한 통합형 채권등급 평가모형 구축에 관한 연구

박기남* · 이훈영** · 박상국***

A Hybrid Credit Rating System using Rough Set Theory

Ki-Nam Park* · Hoon-Young Lee** · Sang-Kuk Park***

■ Abstract ■

Many different statistical and artificial intelligent techniques have been applied to improve the predictability of credit rating. Hybrid models and systems have also been developed by effectively combining different modeling processes or combining the outcomes of individual models.

In this paper, we introduced the rough set theory and developed a hybrid credit rating system that combines individual outcomes in terms of rough set theory. An experiment was conducted to compare the prediction capability of the system with those of other methods. The proposed system based on rough set method outperformed the others.

1. 서 론

IMF이후 기업의 부실채권 처리문제가 금융기관들의 현안으로 등장하였다. 따라서 어느 때보다 기업의 신용을 합리적이고 객관적으로 평가하고 그 결과에 따라 대출여부를 결정하는 과학적인 기업

신용평가 시스템 구축이 절실한 시점이다. 기업신용평가는 기업의 재무변수 또는 비재무변수를 이용하여 기업의 부실위험이나 채권등급을 평가하는 작업이다.

이를 위해 전통적으로 통계적 방법인 다변량 판별분석(multivariate discriminant analysis)이 주로

* 경인여자대학 멀티미디어 정보전산학부 전임강사

** 경희대학교 경영학부 부교수

*** (주) Flexdeal 대표이사

이용되었다. 그러나 판별분석은 사용되는 독립변수가 정규분포를 이루어야 한다는 엄격한 통계적 가정을 만족시켜야 하며 정확한 판별점을 구하기 어렵다는 단점이 지적되어 왔다(Ohison, 1980; 정혜영, 1991). 이러한 판별분석의 단점을 극복하고자 1980년대 후반부터 AI(artificial intelligence) 분야에서 인공신경망(artificial neural network), 사례기반추론, ID3, CART, 유전자 알고리즘등과 같은 방법론에 대한 연구가 이루어지면서 다양한 관점에서 개별방법론에 대한 연구가 이루어지고 있다.

특히, AI기반의 방법론은 통계적 방법론에 비해서 독립변수가 정규분포를 해야한다던가 하는 등의 통계적 가정이 필요 없고 그 동안의 연구를 통해서 통계적 방법론에 비해 모형의 성능이 우수하다는 결과를 얻어왔다.

인공신경망 모형은 판별분석이 가지고 있는 문제점들을 해결할 수 있는 장점이 있으나 모형에 대한 이해가 쉽지 않고 모형의 추정에 많은 데이터를 확보해야 한다는 문제를 가지고 있다. 따라서 현재의 연구방향은 평가 모형들간의 장점을 취하고 단점은 보완하는 하이브리드(hybrid)형태의 모형개발을 대안으로 제시한다.

통합모형은 각 단일모형이 가지고 있는 장점과 특성을 수용하여 개별모형보다 안정성 및 성능의 향상을 얻고자 하는 데 그 목적이 있다. 단일모형들이 갖는 단점을 보완하기 위한 통합모형의 개발방향은 크게 두 가지로 요약된다.

첫째 단일모형을 사용하되 결과값을 산출하기 이전에 여러 모형을 결합하여 더 나은 결과를 얻고자 하는 과정통합방법, 둘째 개별모형들이 산출한 값들을 특정한 방법을 이용하여 통합된 하나의 결과값으로 최종 산출하는 결과통합방법이 그것이다. 그러나 통합모형은 모형들간의 결합에 대한 논리적인 기준의 마련이 어렵고 각 모형들간의 결과값이 상충되는 경우에 이를 해결해야 하는 문제가 뒤따른다.

본 연구는 개별 모형들의 장점을 살린 결과값을 산출하고 각 모형들이 산출한 결과값을 바탕으로

모형을 통합하는 결과통합방식을 따른다. 지금까지 채권등급 평가에 대표적으로 적용된 통계적 모형인 판별분석과 인공신경망 모형을 러프집합을 이용하여 결합하는 통합방법론이 모형의 결과값들간의 상충문제를 자연스럽게 해결하고 그 성과에 있어서도 기존의 제시된 통합방법론보다 더 우수한 결과를 산출하는지 비교분석 하였다.

2. 채권등급평가 모형에 대한 기존 연구

2.1 기존 연구

채권등급 평가는 다중그룹(multi-group)분류문제이므로 기업도산예측과 같은 2(binary)-그룹 분류문제에 비해서 고도화된 전문가 지식이 필요하다(이건창 등, 1996) 따라서 모형설계가 도산예측에 비해서 어렵다

그 동안 채권등급평가 분야에서 연구된 결과를 살펴보면 Fisher가 최초로 회귀분석을 이용하여 채권등급 평가를 한 이후로 판별분석, 로짓분석 등의 통계적 방법론이 채권등급평가에 적극적으로 활용되었다.

그러나 1980년대 후반부터 통계적인 방법론 외에 다양한 인공지능(artificial intelligence)기법들이 소개되면서 채권등급평가 문제에 활발하게 적용되기 시작했고 많은 연구들이 기존방법론에 비해 향상된 결과를 보고해왔다.

Dutta & Shekhar(1988)은 채권등급 평가에 인공신경망을 처음 적용하였고 Kim, Weistroffer & Redmond(1993)는 인공신경망, ID3, 판별분석, 로짓분석을 6개 그룹 채권등급 평가에 적용하여 여러 기법들 중에서 인공신경망의 성능이 가장 높다는 결론을 얻었다.

Surkan & Singleton(1990, 1995)은 판별분석과 인공신경망을 비교하여 판별분석에 비해 인공신경망의 예측정확도가 더 높다는 결과를 발표하였다.

한인구, 권영식, 이진창(1995b)은 인공신경망의

다중그룹 분류문제에 대한 인공신경망의 판별력을 향상시키기 위해 새로운 학습방법인 OPP(ordinal pairwise partitioning)방법을 사용하여 기존 모형의 성능을 향상시켰다. 이견창, 한인구, 김명종(1996)의 연구는 판별분석과 인공신경망 모형의 결과가 다를 경우 이를 다시 인공신경망으로 학습시키는 통합방법론을 제시하고 이 통합방법이 기존의 단일모형보다 더 높은 예측정확도를 보인다는 실증결과를 발표하였다. 신경식(1998)은 유전자 알고리즘(genetic algorithms)과 사례기반추론(case-based reasoning)을 결합한 통합방법론을 제시하면서 기존 모형에 비해서 제시한 통합방법론의 예측력이 크게 개선됨을 보였다. 지금까지 연구된 채권등급 평가에 대한 연구에 대한 요약은 <표 1>과 같다.

<표 1> 기업신용평가 기존연구

저 자	년도	방 법 론
Fisher	1959	Regression Analysis
Pinches, Mingo	1973	판별분석
Kaplan, Urwitz	1979	Probit
Enderington	1985	Multi-nominal Logit
Dutta, Shekhar	1988	ANN*
Surkan, Singleton	1990, 1995	ANN
Kim, Weistroffer, Redmond	1993	판별분석, ID3, Logit, ANN
Moddy, Utans	1995	ANN
한인구, 권영식, 이견창	1995	ANN, OPP**
이견창, 한인구, 김명종	1996	판별분석, ANN, OPP
신경식	1998	CBR, GA***

*Artificial Neural Network **Ordinal Pairwise Partitioning ***Genetic Algorithms

3. 기존 연구 방법론

3.1 판별분석(Multivariable Discriminant Analysis)

판별 분석의 주요 목적은 여러 집단의 중복공간을 최소화 하는 각 집단의 구분선의 선형결합(linear combination) 판별함수를 유도해내는 것이다. 분류되어 있는 집단간의 차이를 설명해줄 수

있는 독립변수들을 찾아내고 이들의 선형결합을 통해 판별식을 만든다. 분류하고자 하는 각 대상들의 속성을 이 판별식에 대입하여 각 대상들이 어느 집단에 속하는지를 판단하는 것이 판별분석이다 따라서 판별 분석의 판별식은 다음과 같은 형태를 갖는다.

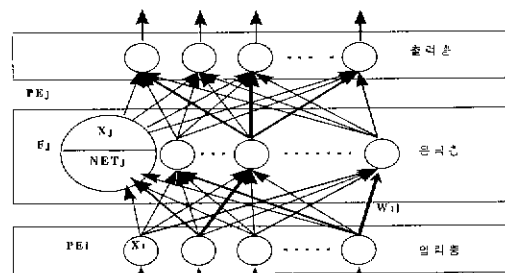
$$Z = W_1X_1 + W_2X_2 + \dots + W_nX_n$$

식에서 Z는 판별점수, W_i 는 i 번째 가중치를 X_i 는 i 번째 독립변수를 나타낸다. 판별식은 각집단 분포가 정규분포를 따른다는 가정하에 유도되므로 정규분포에 대한 가정을 위반하였을 경우 모형의 신뢰도가 떨어진다.

또한 집단간 공분산 행렬이 동일하다고 가정하기 때문에 독립변수가 정규분포에서 벗어나면 모형의 신뢰도가 떨어진다. 이러한 판별함수가 가지고 있는 엄격한 통계적 가정을 가지고 판별식을 구했다하더라도 각 집단을 구분하는 판별점이 명확하지 않다는 한계점은 그대로 남는다.

3.2 인공신경망(Artificial Neural Network)

본 연구에서 로짓모형의 성능과 안전성을 비교하기 위하여 인공신경망을 사용하고 그 결과를 비교하였다. 인공신경망은 생물학적 뇌에 있어서 뉴런(neuron)에 해당하는 처리요소(processing elements ; PE)로 구성된다. 처리 요소는 입력, 순입력, 출력, 가중치, 전이함수 등 5개의 기본요소로 이루어진다. [그림 1]은 처리요소의 기본 구조를 나타낸다.



[그림 1] 인공 신경망 처리 요소의 기본 구조

[그림 1]에서 PE 는 처리 요소, X_i 는 PE_i 의 입력값이 되며, X_j 는 PE_i 와 PE_j 의 출력값이 된다. 그리고 W_{ij} 는 PE_i 와 PE_j 간의 연결강도를 나타내는 가중치이며, net_i 는 순 입력값, f_i 는 값을 X_i 값으로 바꾸어 주는 전이함수이다. 순 입력값과 출력값을 계산해주는 식은 다음과 같다.

$$net_j = \sum W_{ij} X_i, X_j = f_j(net_j)$$

이러한 전이함수에는 계단함수(hard limit), 임계논리함수(threshold), 그리고 S자 형태의 시그모이드 함수(sigmoid)등이 있으며, 그 중에서 가장 많이 쓰이는 것이 시그모이드 함수이다. 또한 동시에 병렬적으로 동작하는 처리 요소들이 모여서 하나의 층(layer)을 이루게 되는데 입력층, 은닉층, 출력층이 있다. 재무분야 연구에서 주로 사용되는 인공신경망 모형은 다층 퍼셉트론(multi-layer perceptron)과 역전파(backpropagation network: BPN) 모형이다(조흥규, 1994).

다층 퍼셉트론은 입력층과 출력층 사이에 여러 개의 은닉층을 두는 것으로 처리요소들 사이의 층간 연결은 전방향 연결형태이며 층내 연결은 없다. 역전파 알고리즘은 입력층의 각 유니트에 입력 패턴을 주면, 이 신호는 입력층의 각 처리 요소에서 변환되어 중간층에 전달되고 최후에 출력층에서 신호를 출력하게 된다. 이 출력값과 목표 출력값 사이의 오차들의 제곱합을 최소로 하기 위해 최급하강법(gradient-descent method)을 사용하며 전방향 계산과 역방향 계산의 두 단계로 이루어진다.

전방향 계산은 주어진 입력값에 대해 신경회로망의 출력값을 계산하는 과정이며, 역방향 계산은 전방향 계산에 의해 얻어진 출력값과 목표출력의 차이, 즉, 오류를 출력층으로부터 입력층으로 역전파 시켜가며 연결가중치들을 조정하고 저장하는 과정을 의미한다.

3.3 방법론들의 특징

본 연구에서 비교하고자 하는 단일모형은 최적

화 기준이나, 절차, 가정, 지식의 표현형태, 지식의 구조등에서 각기 서로 다른 접근방법을 가지고 있다(김광용, 1998). 본 연구에서 통합모형 구성을 이루는 단일모형으로 선정된 판별분석과 인공신경망의 경우를 비교하면 <표 2>와 같다.

<표 2> 판별분석과 인공신경망의 비교

방법론	MDA	ANN
최적화 기준	공분산 행렬	델타 규칙(Delta Rule)
최적화 절차	행렬을 이용한 통계적 추정	역전파 학습
모델의 가정	여러 가지 통계적 가정	없음
지식표현 형태	선형결합 모형	가중치를 가진 네트워크의 연결구조
지식의 구조	반 명제형 지식	비 명제형 지식
주요 장점	안전성(Robustness)	높은 예측력, 안전성
주요 단점	가정의 만족이 어려움	설명력의 부재 (Black Box)

(김광용 1998)

판별분석과 같은 통계적 모형은 반명제형 지식(pseudopositional knowledge)구조를 가지고 있다. 즉, 의사결정에 요구되는 변수들의 중요도에 따른 가중치와 그에 대응하는 변수들의 선형적인 결합을 통해서 모형을 구성하는 부분은 명제형 지식구조에 가깝지만 변수의 통계적 추정은 비명제형 지식(non-propositional knowledge)에 가깝다.

인공신경망의 경우 지식은 추론(inference)을 통해서 얻어지는 것이 아니라 패턴인식(pattern recognition)에 의해서 이루어진다. 다시 말해서 인공신경망은 지식의 표현형태가 여러 노드들간의 가중치가 부여된 네트워크 형태를 이루게 된다.

이러한 인공신경망의 특징으로 인해서 현실세계에서 다루어지는 지식의 표현형태와는 전혀 다른 표현형태를 가지는데 이러한 문제로 인해서 인공신경망에서는 지식을 해석하거나 분석하는 것이 거의 불가능하다. 그러나, 기존의 많은 연구들은 인공신경망이 타 모형에 비해서 일반적으로 높은 예측력과, 안전성 등이 뛰어나다는 것을 보여주고 있다.

그러나 각 모형들이 가지고 있는 독특한 특성으

로 인해서 단일 모형의 경우 편의가 발생할 가능성이 존재한다. 즉, 각 모형이 가정하고 있는 조건에 부합되는 데이터가 주어질 경우에는 예측력이 높아지지만 가정을 만족시키지 못할 경우 예측력이 떨어지게 된다. 따라서 현실적으로 의미있는 모형이 되기 위해서는 각 모형의 특성으로 인한 영향을 최소화시키면서 전반적인 성능의 향상과 안전성을 확보해야 한다.

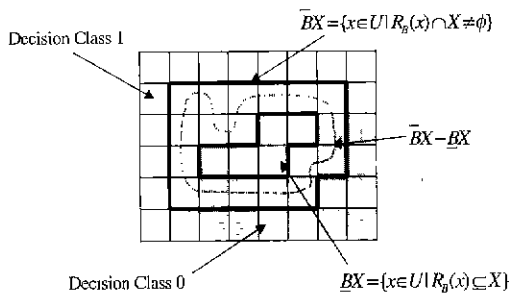
4. 러프집합을 이용한 통합 알고리즘

4.1 러프집합의 개요

러프집합이론은 1970년대에 Zdzislaw Pawlak 교수에 의해서 불확실성 하에서의 의사결정 문제를 해결하기 위한 방법론으로 개발되었다. 다시 말하면 의사결정을 내리기에 불충분한 정보가 있을 경우 이를 해결하기 위한 수학적 방법론이다 (Yao,1996 ; Ohn,1998).

집합의 소속정도가 소속함수(membership function)에 의해서 주어지는 퍼지집합과는 달리 crisp set에 기초한 방법론인데 퍼지집합에서는 일반적으로 집합을 구성하는데 있어 집합의 원소의 포함 유무자체가 확실하지 않다(Ohn,1997).

반면 러프집합에서는 원소 자체는 일반적인 집합구성 요건에 부합되지만 집합을 구성하기 위한 경계영역(boundary region)이 존재한다. 따라서 현실세계에서의 관심영역을 근사공간(approximation space)으로 표현하며 이는 [그림 2]와 같이 upper



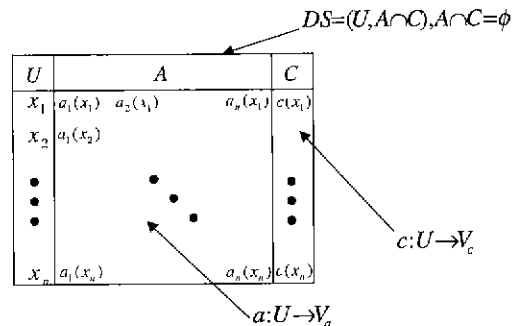
[그림 2] 러프집합

& lower approximation, boundary region의 3가지 구조로 되어 있다. [그림 2]에서 U 는 전체집합 X 는 실제 데이터이고 $R_B(x)$ 는 $x \in X$ 와 같은 패턴을 가지는 원소들의 집합이다.

현실세계에서 다루어지는 데이터의 형태는 [그림 2]와 같은 2차원 표(2 dimensional table)형태의 데이터이다. 이를 러프집합에서는 Decision System으로 새롭게 정의하는데 이것은 2차원 테이블 형태의 데이터로 통계적 개념에서의 독립변수와 종속변수 그리고 관찰값들로 구성되어 있다. 이에 대한 정의는 아래의 수식과 같다.

$$DS = (U, A \cup C), A \cap C = \emptyset$$

여기서 U 는 관찰값들의 전체집합, A 는 조건속성 집합, C 는 decision 속성 집합이라 한다. 이때 임의의 $a \in A$ 에 대해서 $a: U \rightarrow V_a$ 가 주어질 때 V_a 를 a 속성에 대한 value 집합이라 하고 마찬가지로 $c \in C$ 에 대해서 $c: U \rightarrow V_c$ 가 주어질 때 V_c 를 c 속성에 대한 value 집합이라 정의한다. 결국 Decision System은 개념적으로 [그림 3]과 같이 표현된다.



[그림 3] Decision System

의사결정 문제는 [그림 3]에서와 같이 각각의 cell들이 decision 속성의 value set에 의해서 여러 영역으로 나뉘어 지는데 의사결정을 내려야 하는 대부분의 실제 데이터는 여러 영역에 걸쳐서 나타난다.

결국 러프집합에서의 의사결정 문제는 실제 데이터가 여러 영역에 속하지 않으면서 어느 특정한 영역(decision class)에 속하게끔 만들어 주는 새로운 패턴들의 조합을 만들 수 있는가 하는 것과 분류기준이 되는 각각의 규칙에 대해서 확신도를 어떻게 평가하는가 하는 문제로 귀결된다(Ohm, 1998; Yao & Lin, 1996).

4.2 규칙 생성

분류 작업을 위해서 Decision System으로부터 IF-THEN형태의 규칙의 집합을 생성하는데 이때 임의의 $x \in U$ 의 decision 속성이 c 일 경우 이에 대한 규칙은 다음과 같이 정의된다

$$RUL(A, x, c) = \{F_B(x) \rightarrow (c = c(x)) : B \in RED(A, x, c)\}$$

$$F_B(x) = \prod_{b \in B} (b = b(x)), B \subset A.$$

A 는 U 의 전체 condition 속성 집합

따라서 조건 속성(condition attribute)들이 많거나 value set들의 원소가 많은 경우에는 규칙 집합의 규모가 기하 급수적으로 커지는데 이는 분류과정에서 규칙들을 검색하는데 탐색비용이 증가하게 된다는 것을 의미한다.

따라서 각각의 조건 속성들 중 의사결정에 있어서 직접적으로 필요하지 않는 속성들을 삭제하는 과정을 거치는데 이를 Reduction이라 하며 $RED(A, x, c)$ 라 정의하며 이는 x 를 다른 원소와 구별하기 위한 최소한의 속성 집합이다. 이를 한 원소가 아닌 집합 내 모든 원소들로 일반화 시키면 다음과 같다.

$$RUL(A, c) = \bigcup_{x \in U} RUL(A, x, c)$$

위의 공식에 의해서 $x \in U$ 에 대한 규칙이 생성이 되면 이것은 U 의 각각의 원소에 대한 규칙들의 합집합으로 표현된다. 이러한 일반화된 규칙 집합이 구해지면 새로운 데이터에 대한 의사결정은

규칙들의 집합에서 필요한 규칙들을 선택하는 문제가 남는다.

4.3 분류(Classification) 알고리즘

러프집합을 이용한 분류작업은 예측용 데이터에서 생성된 규칙들의 집합에서 검증용 데이터의 각각의 관찰값에 대응(match)되는 규칙을 선택하는 작업이다(Ohm, 1999). 러프집합에서 분류를 하는데 있어서 사용되는 알고리즘은 특별하게 정해져 있지 않다

모형 결합문제에 있어서 가장 어려운 점은 각 모형들에 의해서 나온 결과들이 서로 상충되었을 때 이를 적절하게 해결하는 문제인데 이를 해결하기 위해서 사용한 알고리즘은 러프집합 분석도구인 RosettaTM에서 제공하는 Standard Voter 알고리즘이다. Standard Voter 알고리즘의 분류절차는 다음과 같다.

STEP 1

관찰값 $y \in T$ 의 $F_b(y) \rightarrow (c = c(y))$ 가 $\gamma \in RUL(A, c)$ 인 γ 가 존재하면 $y \in T$ 는 γ 에 의해 분류된다.

STEP 2

관찰값 $y \in T$ 의 $F_b(y) \rightarrow (c = c(y))$ 가 $\gamma \in RUL(A, c)$ 인 γ 를 발견할 수 없다면 $y \in T$ 는 오분류(misclassified)로 처리한다.

STEP 3

관찰값 $y \in T$ 의 $F_b(y) \rightarrow (c = c(y))$ 가 복수개 존재할 경우 $\gamma \in RUL(A, c)$ 인 γ 를 선정할 때 $certainty(x, \beta)$ 가 높은 γ 를 선택한다.

Standard Voter 알고리즘에서 각 모형이 산출한 결과값이 서로 상충될 경우(conflict) 규칙에 대한 확신도 $certainty(x, \beta)$ 를 측정하여 의사결정에 필요한 규칙을 결정한다. 확신도에 대한 정의는 다음과 같다.

$$R_\beta = \{\gamma \in RUL(x) : \gamma \text{ predicts } \beta\}$$

$$votes(\beta) = \sum_{r \in R_\beta} votes(r)$$

$$certainty(x, \beta) = \frac{votes(\beta)}{norm(x)}$$

위 공식에서 확신도는 어떠한 규칙 r 이 있다면 이 규칙에 의해서 예측된 결과 β 가 있을 경우 이 규칙에 의해서 지원되는 $x \in U$ 의 개수는 $votes(\beta)$ 로 표현하였고, 규칙 r 를 가지는 $x \in U$ 개수는 $norm(x)$ 로 표현하였다.

5. 통합모형의 설계

5.1 자료구성 및 변수선정

본 연구에서 사용된 자료는 한국신용정보, 한국 기업평가, 한국신용평가 3개사의 1991년도부터 1995년도의 기업신용평가 자료에 해당하는 기업을 표본으로 하였다. 자료에 대한 요약은 <표 3>과 같으며 표에서 1, 2, 3, 4, 5등급은 A+, A, B, C, D등급을 나타내며 자료 중 변수 값이 누락되어 있는 것은 제외하였다. 이러한 결합모형의 성과를 검증하기 위하여 1991년부터 1995년까지의 기업 신용평가 자료 3822개를 사용하였다.

<표 3> 기업신용평가 자료 통계

등급	1991년	1992년	1993년	1994년	1995년	합계
1	50 (5.6%)	55 (6.8%)	31 (6.2%)	53 (6.6%)	44 (8.9%)	253 (6.6%)
2	187 (20.8%)	182 (22.6%)	152 (18.6%)	181 (22.5%)	108 (21.9%)	810 (21.2%)
3	301 (33.6%)	278 (34.4%)	278 (34.0%)	256 (31.8%)	181 (36.6%)	1294 (33.9%)
4	336 (37.5%)	274 (34.0%)	333 (40.7%)	293 (36.4%)	156 (31.6%)	1392 (36.4%)
5	23 (2.3%)	18 (2.25%)	4 (0.5%)	23 (2.9%)	5 (7.0%)	73 (1.9%)
합계	897 (100%)	807 (100%)	810 (100%)	806 (100%)	494 (100%)	3822 (100%)

본 실험에서는 모형의 성능을 검증하기 위한 학습용 데이터와 검증용 데이터를 구성하는 규칙은 1991년도 데이터를 학습용 데이터로 사용하고

1992년도 데이터를 검증용 데이터로 사용하고 1992년도를 학습용 데이터로 사용할 때에는 1993년도의 데이터를 검증용 데이터로 사용하는 방법으로 모형의 추정과 검증을 반복 수행하였다.

연구를 위해서 선택된 변수는 총 148개의 변수 중 기업심사 전문가와 기존의 신용평가에 관한 문헌연구를 토대로 신용등급간 판별력을 높여 줄 수 있을 것으로 예상되는 변수들과 이찬강, 한인구, 김명중(1996)의 선행연구를 바탕으로 26개의 변수를 선정하였다. 재무제표로부터 추출된 재무변수와 비재무변수로 나눌 수 있는데 규모지표 9개, 수익성 지표 7개, 안전성 지표 5개, 현금흐름지표 4개, 생산성 지표 2개를 선정하였으며 이들 변수에 대한 기초 통계는 <표 4>에 요약되어 있다.

<표 4> 실험에 사용된 변수의 요약

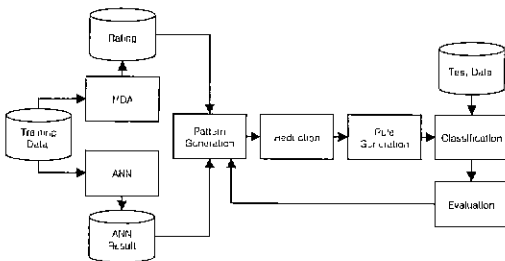
변수 번호	변수이름	평균	표준편차	단위
X1	총자산	261170246	669049369	원
X2	유형고정자산	91695471	272468166	원
X3	자기자본	58457223	160081826	원
X4	매출액	247393440	829115278	원
X5	부가가치	52901480	139914091	원
X6	총차입금	128654693	335598729	원
X7	총수익	256271864	849442823	원
X8	감가상각비	8470428	34183485	원
X9	1인당 매출액	251818	679346	원
X10	업력	22.6582	11.9080	년
X11	총자본사업이익율	9.3708	5.9490	비율
X12	유보액/총자산	12.6599	12.8921	비율
X13	매출액총이익률	19.3151	12.7259	비율
X14	금융비용/매출액	7.0587	6.5284	비율
X15	금융비용/부채	8.2444	3.0294	비율
X16	금융비용/총비용	6.7283	4.2111	비율
X17	감가상각비/총비용	3.4526	3.5051	비율
X18	고장자산구성률	44.8652	18.4426	비율
X19	차입금의존도	46.1144	16.2156	비율
X20	자기자본 비율	23.3110	13.4690	비율
X21	채고자산/유동자산	30.1362	18.1092	비율
X22	단기차입금/총차입금	51.2011	23.9690	비율
X23	현금흐름/총부채	8.1830	8.7868	비율
X24	현금흐름/고정부채	17.3893	29.1981	비율
X25	부채상환계수	42.4134	39.3793	비율
X26	순현금흐름/총자산	4.5841	20.2222	비율
X27	총자본투자효율	28.7165	16.6759	비율

실험자료의 구성을 위한 기초분석과 판별분석 실험을 위해 SAS 6.12를 이용하였고 인공신경망 실험을 위해서 NeuroShell 2를 이용하였다. 인공신경망의 경우 running rate는 0.75, 시스템 에러율은 0.01로 설정하여 실험하였다. 러프집합을 이용한 통합모형의 실험을 위해서는 Rosetta v1.45를 이용하였으며 입력값으로는 각각의 모형에서 산출된 채권등급이 사용되었다.

5.2 통합모형 설계

본 연구에서는 앞절에서 선정된 변수들을 이용하여 판별분석과 인공신경망의 모형을 연도별로 구축하고 검증용 셋을 이용하여 각 모형이 추정하는 기업의 신용등급을 결과값으로 산출하였다. 본 연구에서는 판별분석과 인공신경망에서 얻어진 결과값을 러프집합을 이용하여 결합하는 방법으로 통합모형을 구성하였다. 그러나 이렇게 결합하고자 하는 두 모형의 결과값은 완전히 일치하지 않는다. 일치하는 부분은 문제가 되지 않지만 두 모형의 결과값이 상충이 일어나는 부분은 어느 한쪽의 의견으로 조정해주는 갈등해결이 필요하다.

이를 위해 본 연구에서 제시하는 통합모형의 구조는 [그림 4]와 같다. 본 통합모형은 판별분석과 인공신경망에서 산출된 채권등급 결과를 가지고 패턴을 도출하고 의사결정에 불필요한 속성들을 삭제하는 reduction과정을 거쳐서 규칙집합을 구한다. 이 규칙들을 가지고 분류작업을 수행하는데 모형을 평가하는 과정에서 원하는 신뢰수준에 들지 못하면 다시 패턴을 산출하는 프로세스로 피드백



[그림 4] 통합모형의 구조

과정을 거친다.

본 러프집합을 이용한 통합모형은 각 모형의 산출값을 입력값으로 하여 어떠한 입력패턴을 가질 경우 참인 산출값을 해답으로 줄 수 있는지에 관한 규칙을 확신도를 통하여 구축한 다음 검증용 데이터에 적용하는 방법으로 통합모형의 예측능력을 분석하였다. 또한 본 연구의 벤치마크인 신경망통합방법론은 선행연구에서 사용한 방법을 그대로 반복하였다.

6. 실증분석

6.1 단일모형 실험

본 연구에서 수행한 실험의 결과를 살펴보면 다음 <표 5>와 같다. 먼저 판별분석의 예측정확도는 검증용 셋에서 59.4%에서 61%정도로 나타났으며 이것은 선행연구에서 발표된 것과 대등소이한 결과이다. 각 등급별 평균은 46.5%에서 71.8%정도의 예측정확도를 보였으며 최상위나 최하위등급보다는 중간등급에 있는 표본들의 예측정확도가 대체로 높게 나타났다. 연도별로는 92년의 데이터로 추정하고 93년 데이터 셋으로 검증한 결과가 가장 높은 예측정확도를 보였다.

<표 5> 판별분석실험결과

등급	For92	For93	For94	For95	평균
1	49.1%	52.9%	50.9%	34.1%	46.75%
2	39.6%	61.8%	47.5%	55.6%	51.13%
3	57.6%	62.2%	59.0%	56.4%	58.80%
4	79.2%	63.1%	70.6%	74.4%	71.83%
5	16.7%	25.0%	0.0%	20.0%	15.43%
합계	59.4%	61.7%	58.4%	59.5%	59.75%

인공신경망을 이용하여 수행한 실험의 결과를 살펴보면 <표 6>과 같다. 인공신경망의 예측정확도는 검증용 셋에서 58.2%-65.5%정도의 예측정확도를 나타냈으며 이것은 선행연구에서 발표된 것보다 다소 낮지만 대등소이하다. 각 등급별 평균은 6.25%에서 74.15%정도의 예측정확도를 보였다. 연

도별로는 편별분석의 결과와 마찬가지로 92년의 데이터로 추정하고 93년 데이터 셋으로 검증한 결과가 가장 높은 예측정확도를 보였다.

〈표 6〉 인공지능경망의 실험결과

등급	For92	For93	For94	For95	평균
1	25.5%	47.1%	66.0%	59.1%	49.43%
2	44.0%	61.8%	53.6%	50.0%	52.35%
3	61.5%	62.6%	53.9%	62.4%	60.10%
4	74.8%	73.0%	75.1%	73.7%	74.15%
5	0.0%	25.0%	0.0%	0.0%	6.25%
합계	58.2%	65.5%	60.8%	62.3%	61.70%

6.2 통합모형에 대한 실험

앞의 소절에서 추정된 모형을 이용하여 산출한 결과값을 바탕으로 모형간의 상충관계를 제어하면서 개별 모형을 통합하는 결과통합모형간의 예측능력을 비교한다. 먼저 갈등이 발생한 관측치들을 다시 인공지능경망으로 재학습을 시킴으로서 갈등을 해결하는 또 하나의 모형을 만들고 이러한 모형을 통하여 갈등이 발생한 추정치들이 하나의 추정값을 갖도록 하는 인공지능경망 통합방법론(Integration Method Using ANN; 이후 IMUA로 표기함)의 예측성능을 분석한다.

그 다음으로 본 연구에서 제안하는 러프집합을 이용한 통합방법론(Hybrid Method Using Rough Set Theory; 이후 HMUR로 표기함)의 예측성능을 분석한다. IMUA의 예측결과를 요약한 내용은 <표 7>에 요약되어있다.

〈표 7〉 인공지능경망 통합(IMUA) 통합모형의 예측결과

등급	For92	For93	For94	For95	평균
1	30.9%	35.3%	45.3%	54.5%	41.50%
2	47.8%	62.5%	55.2%	50.0%	53.87%
3	64.0%	64.7%	54.3%	64.1%	61.78%
4	74.1%	73.9%	74.1%	75.6%	74.43%
5	11.1%	25.0%	0.0%	0.0%	9.03%
합계	60.3%	66.0%	60.5%	63.2%	62.75%

역시 통합방법론인 IMUA의 예측정확도는 앞

절에서 제시한 개별 방법론의 결과보다 다소 개선된 예측정확도를 보이고 있다. 두 모형간의 객관적인 성능비교가 가능한 검증용 셋을 기준으로 예측정확도는 60.5%에서 66%정도로 나타났으며 이것은 선행연구에서 발표된 것보다 다소 낮은 예측정확도이나 대등소이한 결과로 볼 수 있다.

또한 각 등급별 평균은 9.02%에서 74.43%정도의 예측정확도를 보였으며 최상위나 최하위등급보다는 중간등급에 있는 표본들의 예측정확도가 대체로 높게 나타났다. 연도별로는 92년의 데이터로 추정하고 93년 데이터 셋으로 검증한 결과가 가장 높은 예측정확도를 보였다. 다음으로 HMUR의 예측정확도를 분석한 결과를 요약하면 <표 8>과 같다.

〈표 8〉 러프집합 통합모형(HMUR) 통합모형의 예측결과

등급	For92	For93	For94	For95	평균
1	43.6%	64.7%	77.4%	86.4%	68.02%
2	64.8%	68.4%	71.3%	76.9%	70.35%
3	59.0%	60.0%	62.9%	79.6%	65.38%
4	74.0%	80.1%	71.3%	66.7%	73.03%
5	44.4%	0.0%	26.1%	0.0%	17.63%
합계	64.0%	69.8%	64.1%	74.7%	68.15%

모든 데이터 셋에서 HMUR의 예측정확도는 IMUA의 예측정확도를 능가하고 있으며 앞에서 분석한 개별 방법론의 어떠한 데이터 셋보다도 더 개선된 결과를 보이고 있다. 개별 데이터 셋을 살펴보면 검증용 셋에서의 예측정확도는 64%에서 74.7%로 나타났고 각 등급별 평균을 살펴보면 17.63%에서 73.03%의 예측정확도를 보였으며 최하위등급보다는 상위등급의 예측정확도가 높게 나타났다. 연도별로는 앞의 결과와는 다르게 94년의 데이터로 추정하고 95년 데이터 셋으로 검증한 결과가 74.7%로 가장 높은 예측정확도를 나타냈다.

6.3 두 통합모형간 예측정확도 비교

앞의 소절에서 제시한 IMUA와 HMUR 두 통합모형의 예측정확도가 실제로 통계적으로 유의한 차이가 있어서 본 연구에서 제시한 HMUR통합모

형이 선행연구에서 제안한 IMUA모형보다 더 예측능력이 낫다고 주장할 수 있는지를 살펴보았다. 예측정확도의 비교는 예측정확도의 비교를 통하여 이루어졌으며 두 통합모형의 예측정확도는 비율 t-test를 통하여 수행되었다 그 결과는 <표 9>에 요약되어 있다.

<표 9> 통합모형의 예측정확도 비교

데이터 셋	검증셋(%)		T-value
	신경망통합	러프셋통합	
FOR92	60.3%	64.0%	-1.701**
FOR93	66.0%	69.8%	-1.657**
FOR94	60.5%	64.1%	-1.513*
FOR95	63.2%	74.7%	-3.936**
평균	62.50%	68.2%	-2.321**

** 는 단측검증시 95%신뢰구간에서 통계적으로 유의함¹⁾
* 는 단측검증시 90%신뢰구간에서 통계적으로 유의함

<표 9>에서 알 수 있듯이 IMUA 통합방법과 HMUR 통합방법의 예측정확도를 비교해 보면 HMUR의 예측정확도가 모든 년도에 걸쳐서 유의하게 우수한 것으로 나타났다. 특히 For95 데이터 셋의 경우에는 HMUR이 IMUA의 결과치 보다 11.5%까지 개선된 결과를 보여 99%신뢰구간에서도 HMUR모형이 더 우수한 것으로 나타났다.

7. 결 론

기업의 신용등급을 평가하는 모형의 개발은 아주 오래된 회계정보시스템에서의 연구 주제였다. 그러나 많은 연구들이 이 연구분야에 새로운 AI기법들을 적용하는 연구였는데 최근에는 각 방법론들간의 통합을 통하여 개별 방법론의 단점을 극복하려는 연구들이 이루어지고 있다.

본 연구는 이러한 맥락에서 이루어진 연구이며 특히 통계적모형과 AI기법들의 개별 결과치를 인정해주면서 이러한 결과치를 통합하여 더 개선된

결과를 얻는 결과치 통합기법으로 러프집합을 이용한 통합모형을 제안하였다. 이러한 러프집합은 모형들의 결과값이 상이해서 나타나는 모형간의 상충문제를 해결할 수 있다.

본 연구에서 제안한 통합방법론의 우수성을 검증하기 위하여 기존 연구에서 사용되었던 인공신경망을 이용한 통합방법론과 비교실험을 수행하였는데 모든 데이터 셋에서 통계적으로 유의한 개선효과를 나타내었다.

본 연구의 한계점으로는 사용된 데이터 셋이 1991년에서 1995년의 기간에 불과하여 너무 짧은 기간을 통한 분석이었다는 점과 오랜시간이 경과된 데이터를 사용하였다는 점과 결과통합을 위한 여러 방법론과 차례로 비교하지 못했다는 것도 한계점으로 남는다.

또한 본 연구에서 제시한 통합모형이 비교적 하위등급보다는 상위등급에서 월등한 예측력의 향상을 보였는데 이는 분류에 필요한 정보량이 하위등급보다는 상위등급에 많았기 때문으로 보인다. 그러나, 실무에서는 상위등급의 예측에 비해서 하위등급에서 예측력의 중요도가 더 높을 수 있다. 따라서 향후 연구에서는 본 모형에서 제시된 통합모형이 이러한 점을 고려하여 하위등급에서의 예측력도 향상시키기 위한 이론적 보완이 함께 이루어져야 할 것이다

참 고 문 헌

- [1] 김광용, "여러가지 Inductive 방법에 대한 통합모델 개발과 그 실증적 유효성에 대한 연구", 「한국경영과학회지」, 제23권 3호(1998.9), pp.185-207.
- [2] 김광용·손광기·홍은선, "Data Mining 기법들과 전문가들로부터 추출된 지식에 관한 실증적 비교연구", 「한국전문가시스템학회지」, 제4권 1호(1998.6), pp.41-58.

1) 본 연구에서 단측검증을 실시한 이유는 본 연구에서 제시한 통합방법론이 기존의 통합방법론보다 우수하다는 가설을 세우고 검증을 하기 때문이다.

- [3] 신경식, “채권등급 평가를 위한 통합형 사례 기반추론 모형 구축,” 박사학위 논문, KAIST, 1998.
- [4] 정혜영, “재무제표정보와 주가수익률 예측,” 『회계학연구』, 제12권, 7월(1991), pp.31-56
- [5] 조홍규, “판별분석. 유사추론, 신경회로망을 이용한 도산예측,” 석사학위논문, KAIST, 1994.
- [6] 한인구, 권영식, 이견창(1995a), “A review of Artificial Intelligence Models in Business classification”, 『한국전문가시스템학회지 창간호』(1995b), pp.23-41.
- [7] 한인구, 권영식, 이견창(1995b), “지능형기업 신용평가시스템의 개발:NICE-AI”, 『경영학연구』, 24권 4호(1995b), pp. 91-117.
- [8] 이견창, 한인구, 김명중, “통계적 모형과 인공 지능 모형을 결합한 기업신용평가모형에 관한 연구,” 『한국경영과학회지』, 21권 1호(1996), pp.81-100.
- [9] Dutta, S and S. Shekhar, “Bond rating:A Non-Conservative Application of Neural network”, *Proceeding of the IEEE International Conference on Neural Networks, II*(1998), pp.433-450.
- [10] Kaplan, R. and G. Urwitz, “Statistical Models of Bond Ratings : A Methodological Inquiry”, *Journal of Business*(1979), pp.231-261.
- [11] Kim, J., Weistroffer and Redmond, “Expert systems for Bond Rating:A Comparative Analysis of Statistical, Rule-based and neural Networks System”, *Expert System, Vol.10, No.3*(1993), pp.167-172. .
- [12] Ohlson, J., “Financial Ratios and the Probabilistic Prediction of Bankruptcy”, *Journal of Accounting Research*, Spring(1980), pp. 109-131.
- [13] Surkan, A., and J. Singleton, “Bond Rating with Neural Network”, In reference, A.(Eds.). *Neural Networks in the Capital Markets*, John Wiley, Chichester(1995).
- [14] Surkan, A. and J. Singleton, “Neural Networks for Bond Rating Improved by Multiple Hidden Layers”, *Proceedings of the IEEE International Conference on, Neural networks*(1990), pp.157-162.
- [15] Øhrn, A., “ROSETTA Technical Reference Manual”, *Norwegian University, Dept. of Computer and Information Science.* (1999).
- [16] Øhrn, A. and J. Komorowski, “ROSETTA : A Rough Set Toolkit for Analysis of Data”, *Proc. Third International Joint Conference on Information Sciences, Fifth International Workshop on Rough Sets and Soft Computing (RSSC'97)*, Durham, NC, USA, March 1-5, Vol.3(1997), pp.403-407.
- [17] Øhrn, A. and J. Komorowski, A. Skowron, P. Synak, “The ROSETTA Software System”, *Bulletin of the International Rough Set Society*, Vol.2, No.1(1998), pp.28-30.
- [18] Yao, Y.Y.(1996). “Two views of the theory of rough sets in finite universes”, *International Journal of Approximation Reasoning*, Vol.15, No.4, pp.291-317.
- [19] Yao, Y.Y. and T.Y Lin, “Generalization of rough sets using modal logic”, *Intelligent Automation and Soft Computing. An International Journal*, Vol.2, No.2(1996), pp.103-120.