

유전 알고리즘 기반 집단분류기법의 개발과 성과평가 : 채권등급 평가를 중심으로

민재형*[†] · 정철우*

Design and Performance Measurement of a Genetic Algorithm-based Group Classification Method : The Case of Bond Rating

Jae H. Min* · Chulwoo Jeong*

■ Abstract ■

The purpose of this paper is to develop a new group classification method based on genetic algorithm and to compare its prediction performance with those of existing methods in the area of bond rating. To serve this purpose, we conduct various experiments with pilot and general models. Specifically, we first conduct experiments employing two pilot models : the one searching for the cluster center of each group and the other one searching for both the cluster center and the attribute weights in order to maximize classification accuracy. The results from the pilot experiments show that the performance of the latter in terms of classification accuracy ratio is higher than that of the former, which provides the rationale of searching for both the cluster center of each group and the attribute weights to improve classification accuracy. With this lesson in mind, we design two generalized models employing genetic algorithm: the one is to maximize the classification accuracy and the other one is to minimize the total misclassification cost. We compare the performance of these two models with those of existing statistical and artificial intelligent models such as MDA, ANN, and Decision Tree, and conclude that the genetic algorithm-based group classification method that we propose in this paper significantly outperforms the other methods in respect of classification accuracy ratio as well as misclassification cost.

Keyword : Bond Rating, Genetic Algorithm, Group Classification Method, Classification Accuracy Ratio, Misclassification Cost

1. 서론

채권등급은 자금조달을 목적으로 기업이 발행하는 채권의 안정성을 평가하는 지표이며, 채권금리를 산정하는 기준이 된다. 따라서 채권등급은 투자자에게 뿐만 아니라 채권발행을 통해 자금을 조달하는 기업에게도 매우 중요한 정보이다. 또한 여러 연구들에서 채권등급과 주가와의 연관성에 관해 논의하면서 채권등급이 기업의 가치를 설명해 줄 수 있는 중요한 정보로서 작용한다고 주장하고 있다[5, 7, 8].

채권등급의 평가와 관련한 과학적 연구는 초기에는 주로 다변량판별분석(multivariate discriminant analysis : MDA), 로짓분석(logit analysis), 프로빗분석(probit analysis) 등의 통계기법을 이용하여 이루어졌으나, 최근에는 인공신경망(artificial neural network : ANN), 규칙기반시스템(rule-based system : RBS), 사례기반추론(case-based reasoning : CBR) 등을 이용한 연구들도 보고되고 있다.

본 연구는 새로운 채권등급 평가모형으로서 유전 알고리즘에 기반한 그룹분류기법을 제안하고 그 타당성을 실증적으로 검증하는데 목적이 있다. 이를 위해 먼저, 두 가지 그룹핑 모형을 이용한 예비실험을 수행하여 그 결과를 분석하였다. 그리고 예비실험 결과를 근거로 하여 두 가지 일반화된 그룹핑 모형을 설계하였다. 구체적으로, 유전 알고리즘을 이용하여 그룹 중심과 속성별 가중치를 탐색하는 방법을 개발함으로써 기존의 통계기반 또는 인공지능 기반의 모형보다 분류 정확도를 높일 뿐만 아니라 오분류비용을 낮출 수 있는 그룹핑 모형 두 가지를 구축하였다. 다음으로, 구축된 모형을 이용하여 채권등급의 분류정확도를 구하였으며, 이를 다변량판별분석, 인공신경망, 의사결정나무 등의 기법을 이용한 예측 결과와 비교하였다. 연구 결과, 본 연구에서 제안한 그룹분류기법이 분류 정확도 및 오분류비용 측면 모두에서 기존 기법들보다 우수함을 알 수 있었다.

2. 이론적 배경

채권등급 평가에 관한 초기 연구는 주로 다변량판별분석[3, 13], 로지스틱회귀분석, 프로빗분석[4, 16] 등과 같은 통계적 방법에 기반하였는데, 이 중에서 가장 많이 사용된 모형은 다변량판별분석모형이다. 다변량판별분석모형은 모형의 구현이 간단하고 학습시간도 짧다는 장점이 있지만 독립변수들이 선형성과 정규성 등의 가정을 요구한다는 점에서 한계가 있는데, 실제로 분석에 사용되는 재무비율 자료들은 이 가정에 부합하지 못하는 경우가 대부분이기 때문이다[6, 11, 12].

이러한 이유로 인공신경망이 채권등급 평가를 위한 대안적 기법으로 이용되었다. 인공신경망 모형은 사전에 변수들 사이의 관계를 가정하지 않으므로 적용 범위에 있어 통계적 방법보다 훨씬 넓다고 할 수 있다[2]. 또한 그 분류정확도에 있어서도 여러 연구에서 인공신경망 모형이 기존의 통계적 방법보다 높은 성과를 나타내는 것으로 보고하고 있다[9, 15]. 근래에는 인공신경망 외에도 규칙기반시스템, 사례기반추론[14] 등의 인공지능 기법들을 적용한 연구들도 발표되고 있다.

2.1 유전 알고리즘 기반 그룹분류기법

유전 알고리즘(genetic algorithm : GA)은 생태계의 선택(selection), 교배(cross over), 돌연변이(mutation)와 같은 생물의 유전과 진화 메커니즘을 수리적으로 모형화한 것으로 최적해를 구하기 위해 확률적 검색방법을 이용하여 방대하고 복잡한 공간을 탐색하는 특징을 갖고 있다. 이러한 특징 때문에 유전 알고리즘은 다양한 제약조건을 포함한 문제 상황에서 목적함수를 최적화하는 파라미터 값을 추정하는데 널리 활용되고 있다.

유전 알고리즘을 군집분석에 적용한 최초의 연구는 Bandyopadhyay and Maulik[1]이다. 이들은 데이터에 알맞은 군집의 수와 군집중심을 찾는 데 유전 알고리즘을 이용함으로써 기존의 k-means, fuzzy-

c-means, EM(expectation maximization) 등의 군집분석 방법이 갖고 있는 문제점, 즉, 미리 군집의 수를 정하고 분석을 하는 문제점과 군집중심의 초기값(initial cluster centers)에 따라서 분석결과가 달라지는 문제점을 극복하고자 하였다. 또한 Lin et al.[10]은 유전 알고리즘을 이용하여 분석자료 가운데서 군집중심을 찾음으로써 최적 군집중심의 탐색 시간을 획기적으로 줄이는 방법을 소개하였다.

본 연구에서 제안하는 그룹분류기법은 기본적으로 군집분석(cluster analysis)의 원리를 상당부분 이용한 그룹핑 기법이다. 다만, 기존의 군집분석은 관측값이 속한 군집에 대한 정보 없이 독립변수만을 가지고 분석하여 군집화 하는 방법인 데 반해, 본 연구에서 제안하는 그룹핑 기법은 관측값이 속한 군집을 아는 상태에서 예측군집과 실제군집이 일치하는 비율을 최대화하거나 또는 총오분류비용을 최소화하는 군집화 방법이라는 차이가 있다.

본 연구의 기법에 대하여 구체적으로 논의하기 위해 앞서 우선 군집분석에 대해 간단히 살펴보면 다음과 같다. 군집분석이란 관측값의 특성을 파악하여 비슷한 속성을 가진 관측값끼리 그룹핑하는 기법으로, 동일한 군집에 속해 있는 연구대상들의 속성은 서로 비슷해야 하고, 다른 군집에 속해 있는 연구대상들의 속성은 서로 달라야 한다는 기본적인 원칙을 갖고 있다. 군집분석에서는 이러한 유사성 또는 이질성의 정도를 나타내기 위해 거리의 개념을 이용한다.

군집분석을 위해 가장 흔히 쓰이는 방법으로는 k-means 군집화를 들 수 있는데, 이 방법은 초기에 임의로 만들어진 k 개의 군집을 기초로 최적의 군집을 만들기 위하여 반복적으로 기존의 군집을 재구성하는 과정을 수행한다. 이러한 k-means 군집화에서 나타나는 문제점은 다음과 같다. 첫째, 연구자가 사전에 몇 개의 군집으로 나눌 것인지를 결정해야 한다는 것이다. 둘째, 군집중심의 초기값(initial seed)을 어떻게 선정하느냐에 따라 분석결과는 전혀 다르게 나타날 수 있다. 세 번째 문제는 k-means 군집화뿐만 아니라 fuzzy-c-means, EM 등의 군집

분석에서도 공통적으로 나타나는 문제점으로, 군집화 할 때 해결해야 할 문제에 따라 속성의 가중치는 다를 수 있으나 이를 반영하지 않고 분석을 수행한다는 것이다.

본 연구에서 제안하는 그룹핑 기법은 기본적으로 k-means 군집화의 원리를 이용함과 동시에, k-means 군집화가 갖고 있는 문제점을 극복하기 위한 다음과 같은 두 가지 장치를 제공한다. 첫째는 관측값 사이의 거리를 계산할 때 속성별로 가중치를 차별화하여 적용한다는 것이고, 둘째는 유전 알고리즘을 이용하여 실제그룹과 예측그룹이 일치하는 비율을 최대화하고, 총오분류비용이 최소화 되도록 그룹중심과 가중치를 탐색한다는 것이다.

이러한 장치를 이용하여 그룹핑을 시도하면 다음과 같은 이점을 피할 수 있다. 첫째, 실제그룹의 개수를 예측그룹의 개수로 이용하므로 그룹 수에 대한 연구자의 임의적 판단을 요구하지 않는다. 둘째, 그룹중심을 유전 알고리즘을 이용하여 탐색함으로써 초기값에 따라 분석결과가 달라질 여지가 사라진다. 셋째, 실제그룹과 예측그룹의 일치비율 최대화 및 총오분류비용 최소화라고 하는 목적에 부합하는 속성별 가중치를 구할 수 있다.

2.2 수리적 모형의 구성

본 연구에서 제안하는 그룹핑 모형을 수리적으로 표현하면 다음과 같다. 우선, 그룹중심과 관측값 사이의 동질성 또는 이질성을 나타내는 거리 개념을 표현하면 식 (1)과 같다.

$$d_{ki} = \sum_j (w_{kj} / \sum_j w_{kj}) |X_{kj} - X_{ij}| \quad (1)$$

여기서, d_{ki} : 그룹 k 의 중심과 관측값 i 사이의 거리 ($k=1, 2, 3, 4, 5, i=1, 2, \dots, N$)

X_{kj} : 그룹 k 의 중심에 대한 j 번째 속성의 값 ($j=1, 2, \dots, M$)

X_{ij} : 관측값 i 의 j 번째 속성의 값

w_{kj} : 그룹 k 의 중심과 관측값 간의 거리에 대한 j 번째 속성의 가중치

식 (1)에서 d_{ki} 는 관측값 i 에 대해서 그룹의 개수만큼 계산되어진다. 그 중에서 가장 작은 값에 상응하는 중심의 그룹(k^*)에 관측값 i 를 배속한다.

이렇게 모든 관측값에 대하여 그룹핑하였을 때, 실제그룹($Group(i)$)과 분석을 통해 예측된 그룹(k^*)이 일치하는 비율(H)을 계산한다. 이때 일치도가 최대화되도록 유전 알고리즘을 이용하여 가중치(w_{kj})와 그룹중심(X_{kj})을 찾는다. 이러한 과정을 수식으로 나타낸 것이 식 (2)이다.

$$\begin{aligned} \text{Max } H &= \frac{1}{N} \sum_{i=1}^N CA_i \\ \text{subject to} \\ CA_i &= 1 \quad \text{if } Group(i) = k^*, \forall i \in \{1, 2, \dots, N\} \\ CA_i &= 0 \quad \text{otherwise,} \\ d_{ki}^* &= \min \left(\sum_j (w_{kj} / \sum_j w_{kj}) |X_{kj} - X_{ij}| \right) \end{aligned} \quad (2)$$

여기서, H : 실제그룹과 예측그룹의 적중률
 N : 관측값 수
 CA_i : 관측값 i 의 실제그룹과 예측그룹의 일치여부($i=1, 2, \dots, N$)
 $Group(i)$: 관측값 i 의 실제그룹
 k^* : d_{ki}^* 에 상응하는 그룹($k=1, 2, 3, 4, 5$)
 d_{ki}^* : 그룹 k 의 중심과 관측값 i 사이의 거리 중 최소 거리

다음으로, 식 (2)에 대한 보완적 모형으로 등급별 오분류비용을 차별화하여 오분류비용의 합을 최소화하는 그룹중심과 속성별 거리 가중치를 탐색하도록 한 모형을 식 (3)과 같이 구축하였다.

$$\begin{aligned} \text{Min } TC &= \sum_{i=1}^N C_{ap} \cdot CA_i \\ \text{subject to} \\ CA_i &= 0 \quad \text{if } Group(i) = k^*, \forall i \in \{1, 2, \dots, N\} \\ CA_i &= 1 \quad \text{otherwise,} \\ d_{ki}^* &= \min \left(\sum_j (w_{kj} / \sum_j w_{kj}) |X_{kj} - X_{ij}| \right) \end{aligned} \quad (3)$$

여기서, TC : 총오분류비용

N : 관측값 수

C_{ap} : a 등급을 p 등급으로 예측한 경우의 오분류비용($a=1, 2, 3, 4, 5; p=1, 2, 3, 4, 5$)

CA_i : 관측값 i 의 실제그룹과 예측그룹의 일치여부($i=1, 2, \dots, N$)

$Group(i)$: 관측값 i 의 실제그룹

k^* : d_{ki}^* 에 상응하는 그룹($k=1, 2, 3, 4, 5$)

d_{ki}^* : 그룹 k 의 중심과 관측값 i 간의 거리 중 최소 거리

식 (2)가 등급별 오분류비용이 모두 동일하다는 가정 하에 분류정확도만을 최대화하는 모형인 반면, 식 (3)은 등급별 오분류비용이 오분류 등급의 차이에 따라 다르다고 가정한 모형으로 실무적인 측면에서 보다 현실적인 모형으로 평가할 수 있다.

2.3 본 연구모형과 사례기반추론의 차이점

본 연구에서 제안하는 유전 알고리즘 기반 그룹 분류기법은 예측기업과 실제기업의 유사도를 기준으로 예측결과를 도출하는 방법론이다. 이 점에서 본 연구의 모형은 사례기반추론(CBR: case-based reasoning)과 비슷하다고 할 수 있다. 그러나 본 연구의 모형은 다음과 같은 점에서 사례기반추론과 차별화된다.

첫째, 사례기반추론은 국지적(local) 학습방법인데 비해, 본 연구의 모형은 전체적(global) 학습방법이다. 즉, 사례기반추론은 각 평가대상 기업과 기존 기업 모두를 비교해서 가장 가까운 k 개의 기업을 참조해 예측결과를 생성하는 국지적 학습방법을 이용한다. 반면, 본 연구에서 제시하는 모형은 기존 기업의 속성변수별 최대값과 최소값 사이의 방대한 공간을 탐색하여 예측력을 최대화하는 최적화된 등급별 중심과 평가대상 기업 사이의 거리를 측정하고, 그 거리가 가장 가까운 중심의 등급을 평가대상 기업의 예측등급으로 이용하는 전체적 학습방법을 이용한다.

둘째, 본 연구의 모형은 사례기반추론에 비해 훨씬 짧은 시간 안에 등급예측을 가능하게 한다. 사례기반추론은 속성변수별 가중치를 최적화하기 위해 훈련용 데이터를 두 세트로 나눌 필요가 있는데, 예를 들어, 1,000개의 데이터가 있다고 할 때 이를 6:4 비율로 나누어 600개는 입력사례로, 400개는 가중치를 최적화하기 위한 사례로 이용한다고 가정하자. 그러면 임의의 가중치에 대해서 계산해야 할 거리의 개수는 $600 \times 400 = 240,000$ 개가 된다. 또한 가중치가 최적화되었다고 할 때, 예를 들어, 개별사례가 200개라고 한다면 계산해야 할 거리의 개수는 $600 \times 200 = 120,000$ 개가 된다. 반면, 본 논문에서 제시하는 모형은 가중치를 계산하기 위해서 훈련용 데이터의 각 사례와 등급별 중심점과의 거리만을 계산하면 되므로, 위의 예를 그대로 이용한다고 할 때 임의의 가중치에 대해서 계산해야 할 거리의 개수는 $1,000 \times 5 = 5,000$ 개로 감소하게 된다. 또한 가중치가 최적화되었을 경우에 계산해야 할 거리의 개수도 $5 \times 200 = 1,000$ 개로 감소한다. 따라서 위의 예의 경우, 본 연구의 모형은 사례기반추론과 비교할 때 가중치를 찾기 위해 필요한 거리의 개수는 $1/48$, 최적화된 가중치를 이용하여 등급예측을 위해 계산해야 할 거리의 개수는 $1/120$ 에 불과한 효율적인 방법이다. 그리고 이러한 차이는 가중치의 개수, 평가대상 기업의 개수, 기존 기업의 개수가 증가함에 따라 더욱 현격하게 벌어지게 된다. 또한 이러한 차이로 인해 등급예측을 위해 걸리는 시간도 그만큼 차이가 나게 된다.

3. 실증분석

3.1 분석자료

본 연구에서는 분석을 위해 한국신용평가(주)로부터 제공받은 1991년부터 1995년 사이의 우리나라 기업들의 채권등급 자료를 이용하였다.¹⁾ 모형 구축을 위한 훈련용 자료(3439개)와 이를 검증하기 위한

검증용 자료(393개)로 나누어 분석을 시행하였으며, 자료의 등급별 구성 비율은 <표 1>과 같다.

<표 1> 자료의 등급별 구성 비율

	등급	수	비율(%)
훈련용	A1	228	6.63
	A2	735	21.37
	A3	1160	33.73
	B	1252	36.41
	C	64	1.86
검증용	A1	25	6.36
	A2	84	21.37
	A3	136	34.61
	B	139	35.37
	C	9	2.29

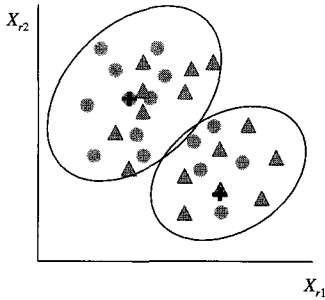
다음으로, 변수선정과 관련하여 본 연구에서는 본 연구의 결과를 기존의 연구결과와 비교할 목적으로 Shin and Han[14]에서 제안한 변수²⁾를 동일하게 사용하였다. 다만, 명목변수에 해당하는 두 변수(상장 및 외감 기업 유형과 기업그룹 순위 유형)는 분석의 편의를 위해 제외하여 본 연구에서는 <표 2>와 같이 총 10개의 변수를 이용하였다.

<표 2> 변수의 정의

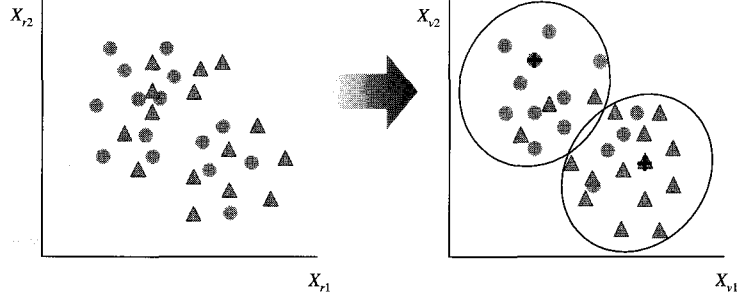
변수	정의
X1	총자산
X2	자기자본
X3	매출액
X4	업력
X5	매출액순이익률
X6	순현금흐름대총자산
X7	재무비용대총자산
X8	차입금의존도
X9	감가상각비대총비용
X10	운전자본회전율

2) Shin and Han[14]은 두 단계에 걸쳐 변수를 선정하였는데, 첫 번째 단계에서는 요인분석, 일원배치 분산분석, Kruskal-Wallis 테스트를 이용하여 27개의 변수(정량적 23개, 정성적 4개)를 선정하였고, 두 번째 단계에서는 첫 번째 단계에서 선정된 27개 변수 중에서 단계별 선택법(stepwise method)을 이용하여 12개의 변수(정량적 10개, 정성적 2개)를 선정하였다.

1) 본 연구에서는 기존 연구와의 성과비교를 위해 Shin and Han[14]에서 이용한 자료를 이용하였다.



[그림 1] [모형 1]의 예



[그림 2] [모형 2]의 예

3.2 연구모형

본 연구에서는 일반화된 모형을 구축하기에 앞서 두 가지 실험적인 모형([모형 1]과 [모형 2])을 만들어 예비분석을 실시하였다. [모형 1]은 모든 속성별 거리 가중치(w_{kj})를 동일한 수준으로 하고, 그룹중심만을 변수로 두어 탐색하도록 한 모형이다. 이 모형에서 그룹중심은 각각의 실제그룹(actual group)에 속한 자료 중에서 찾도록 하였는데, 이는 최적해의 탐색시간을 줄일 수 있을 뿐만 아니라[10], [모형 1]의 결과를 [모형 2]의 결과와 비교하기 용이하기 때문이다. [모형 2]는 속성별 거리 가중치(w_{kj})와 그룹중심(X_{kj})을 동시에 탐색하도록 하는 모형인데, 속성별 거리 가중치(w_{kj})는 0부터 2사이의 범위에서 찾도록 하고, 그룹중심은 [모형 1]에서와 마찬가지로 각각의 실제그룹에 속한 자료 중에서 찾도록 한다. [모형 1]과 [모형 2]의 예를 각각 이차원 평면상에 도시하면 [그림 1] 및 [그림 2]와 같이 나타낼 수 있다.

[그림 1]은 각 관측값의 속성이 X_1 과 X_2 만으로 구성되었다고 가정할 때의 [모형 1]을 나타낸 것이다. '+'로 표시된 두 개의 그룹중심은 관측값 중에서 채택되고, 각 그룹중심으로부터 거리가 가까운 관측값들이 같은 그룹으로 분류된다. 이때, 각 그룹중심은 관측값 중에서 실제그룹과 예측그룹의 일치도를 최대화하는 값으로 선택된다.

한편, [그림 2]는 [그림 1]과 마찬가지로 각 관측값의 속성이 X_1 과 X_2 만으로 구성되었다고 가정할 때의 [모형 2]를 나타낸 것이다. X_{1j} 과 X_{2j} 로 이루

어진 실제공간(real space)에 가중치 w_1 과 w_2 를 곱하여 X_{v1} 과 X_{v2} 의 가상공간(virtual space)을 만들고, 그 위에 관측값들을 위치시킨 후, 관측값 중에서 채택된 두 개의 그룹중심으로부터 거리가 가까운 관측값들을 같은 그룹으로 분류한다. 이때, 각 그룹중심과 가중치 w_1 과 w_2 는 실제그룹과 예측그룹의 일치도를 최대화하는 값으로 구해진다.

이러한 두 가지 예비모형의 실험결과를 통해 본 연구에서는 다음과 같은 두 가지 점을 확인하고자 하였다. 첫째는 속성별 거리에 대하여 가중치를 부여하는 것이 예측정확도에 있어 긍정적으로 작용할 것인지의 여부이다. 둘째는 속성별 거리에 대한 가중치를 부여함에 따라 군집중심이 변하는지의 여부이다.

다음으로 두 가지 예비모형의 실험결과를 토대로 일반화 모형인 [모형 3]과 [모형 4]를 구축하였다. [모형 3]은 속성별 거리 가중치(w_{kj})와 그룹중심(X_{kj})을 동시에 탐색하도록 하는 모형으로, 속성별 거리 가중치(w_{kj})는 0부터 2사이의 실수 범위에서 찾도록 하고³⁾, 최적 그룹중심은 실제그룹에 속한 자료들의 속성별 최대값과 최소값을 구하여 그 범위 내에서 탐색하도록 하였다.

여기서 [모형 2]와 [모형 3]의 차이점은 등급별 중심을 정하는데 있어 [모형 2]는 훈련용 자료의 관측값 중에서 등급별 중심을 찾는 반면, [모형 3]은

3) 속성별 가중치가 [0, 100] 범위의 정수로 취해지도록 설정한 추가적 실험결과에 대한 논의는 4.2절에 기술하였다.

훈련용 자료의 변수별 최소값과 최대값 사이의 광범위한 공간 중에서 등급별 중심을 찾도록 설계되었다는데 있다. 따라서 [모형 2]에 비해 [모형 3]은 탐색시간은 많이 걸리지만 보다 일반화된 해를 구하게 된다.

아울러 [모형 3]과 함께 본 연구에서는 [모형 4]를 추가적으로 제안하였는데, 이 모형은 등급별 오분류비용의 합을 최소화하는 속성별 거리 가중치와 그룹중심을 찾는 모형이다.

한편, 유전 알고리즘을 적용하는데 있어 본 연구 모형이 가지는 염색체(chromosome) 구조를 설명하면 다음과 같다. 각 염색체는 총 60개의 유전자(gene)로 이루어져 있는데, 가중치에 대한 유전자는 10개이고, 등급별 중심에 대한 유전자는 등급별로 10개씩 설정하여 총 50개이다. 가중치는 [0, 2] 범위의 실수이고, 등급별 중심은 각 변수별 훈련용 자료의 최소값과 최대값 범위 내에서 실수값을 취하도록 설정하였다.

최적해를 찾기 위해 population 크기는 100으로 하였고, mutation rate는 $w_{k,j}$ 와 $X_{k,j}$ 에 대해서 모두 0.1로, cross over rate는 $w_{k,j}$ 와 $X_{k,j}$ 에 대해서 모두 0.5로 하였다. Stopping condition은 '수동'으로 설정하였으며 최적해의 수렴 여부를 분석자가 판단하여 탐색을 멈추도록 하였다. 분석도구로는 유전 알고리즘 전문 소프트웨어인 *Evolver* 4.0을 사용하였다.

<표 3>은 본 연구의 각 모형에서 최적해를 찾기 위해 고려한 해의 개수(number of trials)를 정리한

것이다.

<표 3> 모형의 최적해 탐색 시행횟수

모형	시행횟수
1	11,381
2	141,709
3	181,922
4	139,233

4. 분석결과와 논의

4.1 실험결과

<표 4>는 [모형 1]의 결과로 나타난 그룹중심이고, <표 5>는 [모형 2]의 결과로 나타난 그룹중심이다. [모형 1]과 [모형 2]의 그룹중심을 비교하면, 그룹 3과 그룹 5의 중심은 동일하나, 그룹 1, 그룹 2, 그룹 4에 대해서는 중심이 변한 것을 알 수 있다. 이는 실제그룹에 대한 예측그룹의 일치비율을 최대화하는 과정에서 속성별 가중치가 변하면서 그룹중심도 변하였기 때문이다.

다음으로, [모형 2]의 속성별 거리에 대한 가중치를 정리한 <표 6>의 $w_{k,j}/\sum w_{k,j}$ 행을 보면, 0.1을 기준으로 하였을 때, X1(총자산), X2(자기자본), X3(매출액), X5(매출액순이익률), X10(운전자본회전율)은 거리에 대한 가중치가 높게 나왔고, 나머지 X4(업력), X6(순현금흐름대총자산), X7(재무비용대총자

<표 4> [모형 1]의 그룹별 중심

그룹 \ 변수	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	0.688	1.445	0.478	0.622	0.248	-0.195	1.23	-1.321	-0.554	0.232
2	-0.262	-0.106	-0.206	0.202	0.874	-0.365	1.874	-1.812	-0.554	0.517
3	-0.148	-0.071	0.157	0.034	-0.134	-0.866	0.417	-1.186	-0.554	-0.064
4	-0.315	-0.333	-0.239	-0.891	-0.21	0.106	-0.999	-0.096	0.692	-1.327
5	-0.306	-0.342	-0.284	-1.144	-13.801	-0.562	-1.267	2.018	0.692	0.612

주) 그룹 1은 A1, 그룹 2는 A2, 그룹 3은 A3, 그룹 4는 B, 그룹 5는 C등급을 나타낸다. 다른 표에서도 마찬가지로 범례가 적용된다.

〈표 5〉 [모형 2]의 그룹별 중심

그룹 \ 변수	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	0.779	1.664	0.689	1.379	0.799	1.658	1.33	-0.827	-0.554	1.226
2	-0.003	0.177	0.091	-0.135	-0.053	-0.458	0.746	-1.87	-0.554	0.09
3	-0.148	-0.071	0.157	0.034	-0.134	-0.866	0.417	-1.186	-0.554	-0.064
4	-0.366	-0.343	-0.268	-0.723	-1	-0.829	-0.073	1.669	1.937	-0.606
5	0.306	-0.342	-0.284	-1.144	-13.801	-0.562	-1.267	2.018	0.692	0.612

〈표 6〉 [모형 2]의 속성별 거리 가중치

가중치 \ 속성	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
w_{kj}	1.5617	1.9510	1.8947	0.0382	1.2382	0.6062	0.5575	0.0918	0.0050	1.0191
$w_{k,j}/\sum w_{k,j}$	0.1742	0.2177	0.2114	0.0043	0.1381	0.0676	0.0622	0.0102	0.0006	0.1137

산), X8(차입금의존도), X9(감가상각비대총비용)는 거리에 대한 가중치가 작게 나왔음을 알 수 있다.

또한 <표 7>을 보면 [모형 1]에 비해 [모형 2]의 예측정확도가 훈련용 자료에 대해서는 7.00%, 검증용 자료에 대해서는 4.84% 향상된 것을 확인할 수 있다.

〈표 7〉 [모형 1]과 [모형 2]의 예측정확도 비교

	훈련용	검증용
모형 1	55.95%	56.74%
모형 2	62.95%	61.58%

이상의 예비모형에 대한 실험 결과를 통해 다음과 같은 두 가지 결론을 내릴 수 있다. 첫째, 속성별 거리 가중치를 일정하게 두고 그룹중심만을 탐색하는 방법보다는 그룹중심과 속성별 거리 가중치를

함께 탐색하는 것이 모형의 예측정확도를 향상시키는데 있어 더 유리하게 작용한다. 둘째, 속성별 거리 가중치를 부여함에 따라 군집중심의 위치는 달라질 수 있다.

한편, <표 8>과 <표 9>는 [모형 3]의 실험결과로서 그룹별 중심과 속성별 거리 가중치를 정리한 것이다. 먼저 <표 8>은 그룹별(등급별) 중심의 속성을 말해주는데, 이는 각 등급의 표준(standard)이 어떠한 속성을 갖는 지에 대한 정보를 제공한다. 그리고 <표 9>는 각 관측값이 어느 그룹에 속하는가를 결정하는데 있어 어떠한 속성이 얼마나 큰 영향을 미치는지에 대한 정보를 제공해준다. 즉, <표 9>를 보면, 채권등급 평가에 있어 X2(자기자본)와 X3(매출액)이 가장 큰 영향을 미치고, X5(매출액순이익률)도 비교적 큰 영향을 미친 것을 알 수 있다.

〈표 8〉 [모형 3]의 그룹별 중심

그룹 \ 변수	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	1.6775	1.2080	1.7545	-0.0107	0.4087	2.6867	1.8093	-1.9895	-0.5755	2.6278
2	4.5337	0.5043	1.2029	-1.0687	-0.1820	-1.8868	3.1862	-1.4119	-0.4796	0.7325
3	0.8581	-0.2738	1.6306	-0.5988	-0.6450	1.8439	-0.5728	-2.0060	0.4713	-0.4024
4	0.1163	-0.3788	-0.3166	-1.1857	-2.3233	0.1306	-0.7933	0.1833	4.1230	-0.8555
5	-1.4146	-0.7642	-0.6541	-2.2091	-3.3575	-1.1003	0.6501	0.4837	0.3977	-1.8751

<표 9> [모형 3]의 속성별 거리 가중치

속성 가중치	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
w_{kj}	0.0358	1.9873	1.9323	0.2310	1.1235	0.1049	0.4052	0.0879	0.4309	0.5410
$w_{k,j}/\sum w_{k,j}$	0.0052	0.2889	0.2809	0.0336	0.1633	0.0152	0.0589	0.0128	0.0626	0.0786

<표 10>은 [모형 3]의 훈련용 자료에 대한 적중률과 검증용 자료에 대한 예측정확도를 요약한 것이다. [모형 3]의 예측정확도는 [모형 2](<표 7> 참조)와 비교했을 때, 훈련용 자료에 대해서는 2.85%, 검증용 자료에 대해서는 2.03% 향상된 것을 확인할 수 있다.

<표 10> [모형 3]의 예측정확도

훈련용	검증용
65.80%	63.61%

<표 11>은 [모형 3]의 예측정확도를 그룹별로 훈련용 자료와 검증용 자료에 대하여 정리한 것이다.

훈련용과 검증용 자료 모두 실제에 대한 예측정확도는 그룹 4가 78.6%와 77.0%로서 가장 높게 나타났다. 그룹 2와 그룹 3은 그 다음의 예측정확도를 보여주고 있는데, 그룹 2의 경우는 훈련용 자료 65.9%, 검증용 자료 61.9%의 예측정확도를 보여주고 있으며, 그룹 3의 경우는 훈련용 자료 59.5%, 검증용 자료 59.6%의 정확도를 보여주고 있다. 한편, 그룹 1의 예측정확도는 다소 떨어지는 수준으로 훈련용 자료의 경우는 44.3%, 검증용 자료는 40.0%가 나왔다. 그러나 그룹 5의 예측정확도는 훈련용 자료 6.3%, 검증용 자료 0%로 나타났는데, 이러한 결과는 모형이 등급별 예측정확도보다는 전체 예측정확도를 최대한하도록 설계되었기 때문으로 판단된다.

<표 11> [모형 3]의 그룹별 예측정확도

실제 \ 예측		예측					합 계
		1	2	3	4	5	
훈련용	1	101 (44.3%)	109 (47.8%)	17 (7.5%)	1 (0.4%)	0 (0.0%)	228 (100%)
	2	33 (4.5%)	484 (65.9%)	199 (27.1%)	19 (2.6%)	0 (0.0%)	735 (100%)
	3	6 (0.5%)	202 (17.4%)	690 (59.5%)	260 (22.4%)	2 (0.2%)	1,160 (100%)
	4	1 (0.1%)	51 (4.1%)	215 (17.2%)	984 (78.6%)	1 (0.1%)	1,252 (100%)
	5	0 (0.0%)	0 (0.0%)	3 (4.7%)	57 (89.1%)	4 (6.3%)	64 (100%)
검증용	1	10 (40.0%)	15 (60.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	25 (100%)
	2	7 (8.3%)	52 (61.9%)	21 (25.0%)	4 (4.8%)	0 (0.0%)	84 (100%)
	3	0 (0.0%)	29 (21.3%)	81 (59.6%)	26 (19.1%)	0 (0.0%)	136 (100%)
	4	0 (0.0%)	1 (0.7%)	31 (22.3%)	107 (77.0%)	0 (0.0%)	139 (100%)
	5	0 (0.0%)	0 (0.0%)	0 (0.0%)	9 (100.0%)	0 (0.0%)	9 (100%)

다음으로, [모형 3]에 대한 보완적 모형으로서 [모형 4]를 구축하였다. [모형 4]는 등급별로 오분류 비용을 차별화하여 오분류비용의 합을 최소화하는 그룹중심과 속성별 거리 가중치를 모색하도록 한 모형이다. [모형 3]이 등급별 오분류비용이 모두 동일하다는 가정 하에 분류정확도만을 최대화하는 모형인 반면, [모형 4]는 등급별 오분류비용을 오분류된 등급 차이에 따라 차별화한 모형으로 실무적인 측면에서 보다 현실적인 모형이다(식 (3) 참조).

본 연구에서는 등급별 오분류비용(C_{op})으로 <표 12>의 자료를 적용하였다. 이는 2006년 7월 10일 기준 3개월 만기 무보증 회사채의 시장수익률을 이용하여 등급간 시장수익률의 차이를 계산한 것으로, 등급별 오분류비용의 대리치로 이용하였다.

[모형 4]의 분석결과는 <표 13>부터 <표 17>에 정리하였다. 구체적으로, <표 13>은 그룹별 중심을, <표 14>는 속성별 거리 가중치를 나타내고 있는

며, [모형 4]의 예측정확도와 오분류비용은 [모형 3]과 비교하여 <표 15>와 <표 16>에 각각 정리하였다. 그리고 <표 17>은 훈련용 자료와 검증용 자료에 대한 [모형 4]의 그룹별 예측정확도를 나타낸 것이다.

<표 15>를 보면, [모형 4]의 예측정확도는 [모형 3]과 비교하여 훈련용 자료에 대해서는 거의 같은 수준이고, 검증용 자료에 대해서는 2.55% 향상된

<표 12> 그룹(등급)별 오분류비용

예측 실제	1	2	3	4	5
1	0.00	0.12	0.29	0.41	0.57
2	0.12	0.00	0.17	0.29	0.29
3	0.29	0.17	0.00	0.12	0.28
4	0.41	0.29	0.12	0.00	0.16
5	0.57	0.45	0.28	0.16	0.00

<표 13> [모형 4]의 그룹별 중심

변수 그룹	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	1.011	1.256	3.389	0.464	1.186	2.004	2.026	-2.056	-0.447	1.360
2	3.339	0.433	0.173	2.406	0.302	0.937	2.492	-1.613	0.930	0.059
3	0.918	-0.223	0.455	0.215	-0.145	3.288	0.674	-1.797	2.562	-0.555
4	-0.340	-0.374	-0.260	0.916	-1.782	-0.743	0.175	-0.087	2.580	-1.133
5	0.126	-0.288	-0.064	1.019	-3.644	3.253	-1.783	-0.257	1.526	0.101

<표 14> [모형 4]의 속성별 거리 가중치

속성 가중치	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
w_{kj}	0.6709	1.7958	1.2530	0.0570	1.0283	0.0408	0.2035	0.0401	1.1273	0.5410
$w_{kj}/\sum w_{kj}$	0.0993	0.2657	0.1854	0.0084	0.1522	0.0060	0.0301	0.0059	0.1668	0.0801

<표 15> [모형 3]과 [모형 4]의 예측정확도

	훈련용	검증용
모형 3	65.80%	63.61%
모형 4	65.77%	66.16%

<표 16> [모형 3]과 [모형 4]의 오분류비용

	훈련용	검증용
모형 3	191.01	20.63
모형 4	177.21	20.03

<표 17> [모형 4]의 그룹별 예측정확도

실제 \ 예측	예측					합 계	
	1	2	3	4	5		
훈련용	1	105 (46.1%)	104 (45.6%)	18 (7.9%)	1 (0.4%)	0 (0.0%)	228 (100.0%)
	2	27 (3.7%)	428 (58.2%)	266 (36.2%)	13 (1.8%)	1 (0.1%)	735 (100.0%)
	3	1 (0.1%)	136 (11.7%)	787 (67.8%)	223 (19.2%)	13 (1.1%)	1,160 (100.0%)
	4	0 (0.0%)	25 (2.0%)	261 (20.9%)	934 (74.6%)	32 (2.6%)	1,252 (100.0%)
	5	0 (0.0%)	0 (0.0%)	1 (1.6%)	55 (85.9%)	8 (12.5%)	64 (100.0%)
검증용	1	10 (40.0%)	14 (56.0%)	1 (4.0%)	0 (0.0%)	0 (0.0%)	25 (100.0%)
	2	1 (1.2%)	52 (61.9%)	29 (34.5%)	2 (2.4%)	0 (0.0%)	84 (100.0%)
	3	1 (0.7%)	21 (15.4%)	93 (68.4%)	19 (14.0%)	2 (1.5%)	136 (100.0%)
	4	0 (0.0%)	1 (0.7%)	32 (23.0%)	104 (74.8%)	2 (1.4%)	139 (100.0%)
	5	0 (0.0%)	0 (0.0%)	0 (0.0%)	8 (88.9%)	1 (11.1%)	9 (100.0%)

결과를 보여주고 있다.⁴⁾ 반면, <표 16>을 보면, [모형 4]의 오분류비용은 [모형 3]의 오분류비용보다 훈련용 자료에 대해서는 13.8이 낮고, 검증용 자료에 대해서는 0.6이 낮은 일관된 결과를 보임을 알 수 있다. 따라서 실험결과에 따르면, 예측정확도 측면에서는 [모형 3]과 [모형 4] 중에 어느 모형이 더 우수한 모형이라고 판단하기 어렵지만, 오분류비용 측면에서는 [모형 4]가 [모형 3]보다 우수한 모형이라고 평가할 수 있다.

다음으로 그룹별 예측정확도를 보면, [모형 3]의 경우에는 그룹 5의 예측정확도가 훈련용 자료 6.3%, 검증용 자료 0%에 그친데 비해(<표 11> 참조), <표 17>에 나타난 [모형 4]의 경우에는 그룹 5의 예측정확도가 훈련용 자료 12.5%, 검증용 자료는 11.1%로 나타나 해당 그룹의 예측정확도는 향상된

것을 알 수 있다. 또한 그룹 3의 경우에도 [모형 4]의 예측정확도가 [모형 3]보다 훈련용 자료는 8.3%, 검증용 자료는 8.8% 향상된 것을 확인할 수 있다. 그러나 그룹 4의 예측정확도는 [모형 4]가 [모형 3]에 비해 훈련용 자료의 경우에는 4%, 검증용 자료의 경우에는 2.2% 하락했음을 알 수 있다. 이러한 결과는 [모형 3]과 비교하여 [모형 4]에서 그룹 3과 그룹 5에 대한 예측정확도가 높아지면서 그룹 4에 대한 예측정확도는 낮아지는 상쇄현상(trade-off)이 발생했기 때문으로 판단된다.

4.2 추가적 실험

유전 알고리즘을 이용하는 연구의 경우, 일반적으로 최적해를 구하는데 많은 시간이 소요되기 때문에 탐색시간을 줄이기 위한 연구가 추가적으로 진행되기도 한다. 이를 위해 본 연구에서는 다음과 같은 시도를 하였다. 지금까지 본 연구에서는 속성별 가중치가 [0, 2] 범위의 실수값을 취하도록 하고,

4) 훈련용 자료에 대해서는 [모형 3]의 예측정확도가 근소한 차이로 높지만, 검증용 자료에 대해서는 [모형 4]의 예측정확도가 높은 이유는 [모형 3]에서 과적합(overfitting) 문제가 발생하였기 때문으로 판단된다.

이 값을 가중치의 합으로 나누어줌으로써 속성별 최종 가중치는 [0, 1] 범위에서 구해지도록 설정하였다. 이 경우, 속성별 가중치를 구하기 위해서는 [0, 2] 범위의 방대한 실수 공간을 탐색하게 되므로 탐색 시간이 많이 소요된다. 따라서 속성별 가중치가 [0, 100] 범위의 정수로서 취해지도록 설정을 하면 예측정확도는 떨어질 수 있지만 탐색시간은 줄어드는 효과를 볼 수 있다. <표 18>은 속성별 가중치를 [0, 100] 범위의 정수로 설정했을 때의 모형별 예측정확도를 구하여, 이를 기존의 예측정확도와 비교한 것이다.

<표 18> 속성별 가중치 설정에 따른 모형별 예측정확도의 비교

	[0, 2] 범위의 실수		[0, 100] 범위의 정수	
	훈련용	검증용	훈련용	검증용
모형 2	62.95%	61.58%	61.47%	63.61%
모형 3	65.80%	63.61%	64.64%	66.16%
모형 4	65.77%	66.16%	64.15%	62.60%

<표 18>을 보면, 훈련용 자료에 대해서는 [모형 2], [모형 3], [모형 4] 모두에서 속성별 가중치를 [0, 2] 범위의 실수로 설정했을 때의 예측정확도가 높게 나온 것을 알 수 있다. 반면, 검증용 자료에 대해서는 속성별 가중치를 [0, 100] 범위의 정수로 설정했을 때 [모형 2]와 [모형 3]의 예측정확도가 높게 나왔고, [모형 4]의 경우에는 속성별 가중치를 [0, 2] 범위의 실수로 설정했을 때의 예측정확도가 높게 나왔음을 알 수 있다. 따라서 전체적으로 보았을 때 속성별 가중치를 [0, 2] 범위의 실수로 설정한 경우의 예측정확도가 근소한 차이지만 높음을 실험적으로 알 수 있다.

4.3 그룹분류모형의 성과비교

4.3.1 분류정확도

본 연구에서 개발한 분류모형과 기존의 분류모형의 성과를 비교하기 위해 동일한 자료를 가지고 다변량판별분석(MDA), 인공신경망(ANN), 의사결정나무

(Decision Tree)를 이용하여 분석을 실시하였다.

우선 다변량판별분석에 있어 변수선택은 단계별 선택법(stepwise selection method)을 이용하였고, 이때의 F-진입값은 0.05, F-제거값은 0.1로 하였으며, 사전확률은 훈련용 데이터의 등급 구성 비율을 기준으로 하였다.

인공신경망 모형은 입력계층과 출력계층, 그리고 하나의 은닉계층을 가지는 3층 퍼셉트론(three layer perceptron)을 사용하였으며, 은닉층의 노드(node) 수는 입력변수의 수와 같은 수를 사용하였다. 모형의 출력층은 각 등급을 더미변수화하여 5개의 노드로 구성하였다. 이때, 출력노드의 값은 [0, 1]의 범위에 존재하게 된다. 또한 입력변수로는 <표 2>의 10개 변수를 이용하였다. 훈련용 자료로는 본 연구의 모형 구축시 훈련용 자료로 이용한 자료를 이용하였고, 학습 종료를 위한 최적의 조건(stopping condition)을 찾기 위한 테스트용 자료와 학습을 통해 구축된 모형을 검증하기 위한 검증용 자료는 본 연구의 모형 구축시 검증용으로 이용한 자료 393개를 194개와 199개로 나누어 각각 사용하였다.

의사결정나무 모형은 분석 알고리즘으로서 C&RT를 이용하였고, 최대 뿌리마디 아래 깊이는 5로 하였으며, 사전 확률은 훈련용 데이터의 등급 구성 비율을 기준으로 하였다.

이상의 방법으로 구축한 3가지 기존 분류모형의 예측정확도와 본 연구에서 제안한 [모형 3] 및 [모형 4]의 예측정확도를 비교하면 <표 19>와 같다.

<표 19> 분석모형별 예측정확도(%)

모형	등급					전체
	1	2	3	4	5	
MDA	40.0	57.1	60.3	67.6	11.1	59.8
ANN	37.5	38.8	45.5	88.6	0.0	60.3
Decision Tree	56.0	56.0	74.3	66.9	0.0	64.9
모형 3	40.0	61.9	59.6	77.0	0.0	63.6
모형 4	40.0	61.9	68.4	74.8	11.1	66.2

<표 19>를 보면, 본 연구에서 개발한 [모형 4]의

전체 예측정확도가 66.2%로서 가장 높은 수준을 보이고, 의사결정나무(Decision Tree)의 예측정확도는 64.9%로서 두 번째로 높은 수준을 보였으며, 그 다음으로 [모형 3], 인공신경망(ANN), 다변량판별분석(MDA) 순으로 예측정확도를 보이고 있다. 그러나 등급별 예측정확도를 보면 등급별로 높은 예측정확도를 보이는 기법이 다른 것을 알 수 있다. 즉, 그룹 1에 대해서는 의사결정나무가, 그룹 2에 대해서는 [모형 3]과 [모형 4]가, 그룹 3에 대해서는 의사결정나무가, 그룹 4에 대해서는 인공신경망이, 그룹 5에 대해서는 다변량판별분석과 [모형 4]가 가장 높은 예측정확도를 보임을 알 수 있다. 따라서 어느 기법이 우수한 지를 판단하려면 추가적인 분석이 필요함을 알 수 있다. 이를 위해 본 연구에서는 다음과 같이 오분류 등급 차이 관점에서 추가적인 통계분석을 수행하였다.

4.3.2 오분류 등급 차이

채권등급 평가에 있어서 오분류비용은 오분류 등급 차이(오차등급)의 크기에 비례한다. 예를 들어, 1등급을 2등급으로 잘못 분류한 경우보다 1등급을 3등급으로 잘못 분류하였을 때, 더 큰 비용이 발생하게 된다. 그러나 전체 분류정확도만을 가지고 분석모형을 평가하게 되면 이 점은 고려할 수 없게 된다. 따라서 오차등급의 크기를 고려한 성과평가를 위해 본 연구에서는 비모수검정기법인 Wilcoxon 부호순위검정(signed rank test)을 실시하였다. <표 20>은 각 모형별 실제등급과 예측등급 간에 유의한 차이가 존재하는지를 파악하기 위하여 분석모형별로 Wilcoxon 부호순위검정을 실시한 결과이다.

<표 20>을 보면, MDA와 본 연구에서 제안한 [모형 3]과 [모형 4]의 예측등급은 실제등급과 유의

수준 5%에서 통계적으로 유의한 차이를 나타내지 않았다. 반면, 유의수준 5%에서 인공신경망과 의사결정나무의 예측등급은 실제등급과 통계적으로 유의한 차이를 보였다. 의사결정나무의 경우, 전체 분류정확도에서는 [모형 4]에 이어 두 번째로 높은 정확도를 보였으나(<표 19> 참조), Wilcoxon 부호순위검정 결과, 실제등급과 예측등급이 유의한 차이를 보인 것은 |실제등급-예측등급|이 큰 경우가 많았기 때문이다. 반면, [모형 3]은 전체 분류정확도에서는 의사결정나무보다 약간 낮은 정확도를 보였지만, Wilcoxon 부호순위검정 결과, 실제등급과 예측등급이 유의한 차이를 보이지 않은 것은 |실제등급-예측등급|이 큰 경우가 적었기 때문이다. 따라서 [모형 3]은 의사결정나무에 비해 전체 분류정확도 측면에서는 약간 낮은 결과를 보이지만, 오분류 등급 차이 관점에서 보면 의사결정나무보다 우수한 모형이라고 말할 수 있다. 결론적으로, 전체 예측정확도와 오분류 등급 차이를 종합적으로 고려할 때 본 연구에서 제안한 [모형 3]과 [모형 4]가 기존의 분석모형에 비해 우수하다는 결론을 내릴 수 있다.

5. 결 론

본 연구에서는 유전 알고리즘 기반 그룹분류기법을 제안하고, 그 타당성을 실증적으로 검증하였다. 일반화된 모형을 구축하기에 앞서 두 가지 예비모형을 만들어 실험을 수행하고, 그 결과를 분석함으로써 속성별 거리에 대한 가중치를 설정하는 방법에 대한 타당성을 확보할 수 있었다. 또한 속성별 거리 가중치를 설정함에 따라 그룹 중심의 위치도 다소 변하는 것을 확인할 수 있었다. 이러한 예비모

<표 20> Wilcoxon 부호순위검정 결과

	MDA-Actual	ANN-Actual	DT-Actual	모형 3-Actual	모형 4-Actual
통계량 값(Z 값)	-1.619	-2.807	-3.099	-0.628	-0.654
근사유의확률(양측)	0.105	0.005*	0.002*	0.530	0.514

주) * p < 0.05.

형의 실험결과를 토대로 일반화된 모형을 구축하고, 그 성과를 기존의 통계적 모형과 인공지능 모형의 성과와 비교하였다. 연구 결과, 본 연구에서 제안한 유전 알고리즘 기반 그룹분류기법은 분류정확도와 오분류비용 모두를 고려할 때 다변량판별분석, 인공신경망, 의사결정나무 모형에 비해 우수한 기법이라는 결론을 내릴 수 있었다.

본 연구에서 제안한 유전 알고리즘 기반 그룹분류기법은 다음과 같은 장점을 가진다. 첫째, 분석을 통해 그룹별 중심의 속성을 알 수 있는 장점이 있다. 그룹별 중심의 속성은 각 그룹의 표준에 대한 정보를 제공해준다. 둘째, 속성별 거리에 대한 가중치를 통해서 그룹을 결정하는데 있어 영향력이 큰 변수가 무엇인지를 파악할 수 있다. 셋째, 분석자가 통제해야 할 파라미터가 거의 없다는 것이다. 인공신경망이나 의사결정나무 모형의 경우 자료에 따라서 예측정확도를 높이기 위해 분석자가 조절해야 하는 매개변수들이 많은 편이다. 이에 반해 유전 알고리즘 기반 그룹분류기법은 분석자가 조절해야 하는 부분이 거의 없으므로 분석자의 연구능력에 상관없이 비교적 일관성 있는 결과를 얻을 수 있다.

그러나 본 연구에서 제안한 그룹분류모형의 이러한 기여에도 불구하고 본 연구는 다음과 같은 한계점을 갖는다. 첫째, 본 연구에서 제안한 [모형 3]의 경우, 훈련용 자료의 예측정확도를 최대화하도록 설계하였으므로 검증용 자료의 예측정확도가 훈련용 자료의 예측정확도에 비해 심하게 낮게 나타나는 과적합 문제가 발생할 가능성이 있다. 따라서 향후 연구에서는 인공신경망 모형을 구축할 때와 같이 과적합 문제를 방지할 수 있는 장치를 보완할 필요가 있다. 둘째, 본 연구의 모형은 유전 알고리즘을 이용하고 있는데, 유전 알고리즘의 특성상 최적의 그룹중심과 속성별 거리 가중치를 구하기 위한 탐색시간이 다른 방법에 비해 오래 소요되는 편이다. 이는 실무적인 관점에서 단점이라고 할 수 있다. 따라서 본 연구의 방법론이 보다 실용성을 갖추기 위해서는 탐색시간을 줄이기 위한 보완적 방법에 대한 연구가 지속되어야 할 것이다.

참고 문헌

- [1] Bandyopadhyay, S. and U. Maulik, "Genetic Clustering for Automatic Evolution of Clusters and Application to Image Classification," *Pattern Recognition*, Vol.35, No.6 (2002), pp.1197-1208.
- [2] Cheng, B. and D.M. Titterton, "Neural Networks: A Review from a Statistical Perspective," *Statistical Science*, Vol.9, No.1 (1994), pp.2-30.
- [3] Deakin, E.B., "Discriminant Analysis of Predictors of Business Failure," *Journal of Accounting Research*, Vol.10, No.1(1972), pp.167-179.
- [4] Ederington, H.L., "Classification Models and Bond Ratings," *Financial Review*, Vol.20, No.4(1985), pp.237-262.
- [5] Ederington, L.H., J.B. Yawitz, and B.E. Roberts, "The Information Content of Bond Ratings," *Journal of Financial Research*, Vol.10, No.3(1987), pp.211-226.
- [6] Ezzamel, M., J. Brodie, and C. Mar-Molinero, "Financial Patterns of U.K. Manufacturing Companies," *Journal of Business, Finance and Accounting*, Vol.14, No.4(1987), pp.519-565.
- [7] Goh, J.C. and L.H. Ederington, "Is a bond rating downgrading bad news, good news, or no news for stockholders?" *Journal of Finance*, Vol.48, No.5(1993), pp.2001-2008.
- [8] Holthausen, R.W. and R.W. Leftwich, "The Effect of Bond Rating Changes on Common Stock Prices," *Journal of Financial Economics*, Vol.17, No.1(1986), pp.57-90.
- [9] Kim, J., H.R. Weistroffer, and R.T. Redmond, "Expert Systems for Bond Rating: A Comparative Analysis of Statistical, Rule-based and Neural Network Systems," *Expert Sys-*

- tems, Vol.10, No.3(1993), pp.167-172.
- [10] Lin, H.J., F.W. Yang, and Y.T. Kao, "An Efficient GA-based Clustering Technique," *Tamkang Journal of Science and Engineering*, Vol.8, No.2(2005), pp.113-122.
- [11] Maher, J.J. and T.K. Sen, "Predicting Bond Ratings Using Neural Networks : A Comparison with Logistic Regression," *Intelligent Systems in Accounting, Finance and Management*, Vol.6, No.1(1997), pp.59-72.
- [12] McDonald, B. and M.H. Morris, "The Statistical Validity of the Ratio Method in Financial Analysis : An Empirical Examination," *Journal of Business, Finance and Accounting*, Vol.11, No.1(1984), pp.89-97.
- [13] Pinches, G.E. and K.A. Mingo, "A Multivariate Analysis of Industrial Bond Ratings," *Journal of Finance*, Vol.28, No.1(1973), pp. 1-18.
- [14] Shin, K.S. and I.G. Han, "Case-based Reasoning Supported by Genetic Algorithms for Corporate Bond Rating," *Expert Systems with Applications*, Vol.16, No.2(1999), pp. 85-95.
- [15] Singleton, J.C. and A.J. Surkan, "Neural Networks for Bond Rating Improved by Multiple Hidden Layers," *Proceedings of the IEEE International Conference on Neural Networks*, Vol.2(1990), pp.163-168.
- [16] Wiginton, J.C., "A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behavior," *Journal of Financial and Quantitative Analysis*, Vol.15, No.3(1980), pp.757-770.

