

하이브리드 인공지능을 이용한 파산 유형 예측 신경망 모델*

조남옥
이화여자대학교 경영대학
(이메일 : namok.jo@gmail.com)

김현정
이화여자대학교 경영대학
(이메일 : charitas@empal.com)

신경식
이화여자대학교 경영대학
(이메일 : ksshin@ewha.ac.kr)

파산의 예측은 회계 및 금융 분야에서 광범위하게 연구되었습니다. 리스크 관리 측면에서 대출 결정 및 금융 기관의 수익성에 중요한 영향을 미칠 수 있습니다. 많은 연구자들은 보다 강력한 파산 예측 모델을 구축하는 데 집중해 왔습니다. 초기 연구에서는 파산 예측을 위해 다중 판별 분석(MDA) 및 로지 분석과 같은 통계 기법을 주로 사용했습니다. 그러나 많은 연구에서 인공 신경망(ANN), 의사 결정 트리, 사례 기반 추론(CBR) 및 지원 벡터 기계(SVM)와 같은 인공 지능(AI) 접근 방식이 1990년대부터 통계적 방법에는 적용에 엄격한 가정이 있기 때문에 비즈니스 분류 문제가 발생합니다. 기업 파산에 대한 이전 연구에서 많은 연구자들은 재무 비율을 이용한 파산 예측 모델 개발에 중점을 두었습니다. 그러나 특정 유형의 파산을 제안하는 연구는 거의 없습니다.

이전의 파산 예측 모델은 일반적으로 기업이 파산할지 여부를 예측하는 데 관심이 있었습니다. 파산 유형에 대한 대부분의 연구는 이전 문헌을 검토하거나 사례 연구를 수행하는 데 중점을 두었습니다. 이에 본 연구에서는 국내 중소 건설사의 수익성, 안정성, 활동지표 측면에서 특정 유형의 도산과 도산 발생을 예측하기 위해 데이터 마이닝 기법을 활용한 모형을 개발하였다. 따라서 기업은 이를 사전에 예방할 수 있다.

파산 유형 예측을 위해 두 개의 인공 신경망(ANN)을 사용하는 하이브리드 접근 방식을 제안합니다. 첫 번째는 파산 예측을 위해 지도 학습을 사용하는 역전파 신경망(BPN) 모델이고 두 번째는 비지도 학습을 사용하여 파산 데이터를 여러 유형으로 분류하는 자체 조직화 지도(SOM) 모델입니다. 구축된 모델을 기반으로 모델 개발에 활용되지 않은 검증 세트에 BPN 모델을 적용하여 기업의 파산을 예측합니다. 이를 통해 BPN 모델에서 예측한 파산 데이터를 사용하여 특정 유형의 파산을 식별할 수 있습니다. SOM 모델에서 파생된 클러스터의 특성을 해석하기 위해 각 클러스터에 대한 통계적 검정을 통해 선택된 입력 변수의 평균을 계산했습니다. 각 군집은 파산기업의 자료를 통해 분류된 파산유형을 나타내며, 입력변수는 각 군집의 의미를 해석하는데 있어서 재무비율을 나타낸다.

* 이 작업은 한국정부의 연구비 지원(NRF-2013S1A3A2054667)의 지원을 받았습니다.

이 작업은 2012년 이화여자대학교 연구비 지원을 받았습니다.

실험 결과 5가지 유형의 파산 유형은 각각 재무비율에 따라 특성이 다른 것으로 나타났다. 유형 1(심각한 도산)은 클러스터링 결과에 따라 매출에 비해 EBITDA(이자, 세금, 감가상각 전 이익)를 제외하고 재무제표가 열등합니다. 유형 2(안정성 결여)는 유동비율이 낮고, 총자산 대비 자기자본이 낮고, 총자산 대비 총차입금이 높다. 유형 3(활동 부족)은 총 자산 회전율과 고정 자산 회전율이 약간 낮습니다. 유형 4(수익성 부족)는 수익성 지표를 나타내는 총자산 대비 이익잉여금이 낮고 매출 대비 EBITDA가 낮습니다. 유형 5(회수 가능한 파산)는 파산하더라도 다른 유형의 파산에 비해 상대적으로 재무 상태가 양호한 기업을 포함합니다. 이를 바탕으로 신용평가 분야에 종사하는 연구자 및 실무자들은 기업도산 유형에 대한 보다 유용한 정보를 얻을 수 있다.

본 논문에서는 기업의 재무비율을 활용하여 파산 유형을 분류하였다. 파산을 정확히 예측하고 파산의 종류를 의미 있게 분류할 수 있는 입력변수를 선택하는 것이 중요하다. 추가 연구에서 우리는 기업의 규모, 산업 및 연령과 같은 비재무적 요인을 포함할 것입니다. 따라서 질적 요인을 결합하고 전문가의 도메인 지식을 반영하여 파산 유형에 대한 현실적인 클러스터링 결과를 얻을 수 있습니다.

핵심어 : 파산예측, 파산유형분류, 인공지능경망, 역전파신경망, 자기조직화 지도

접수일 : 2015년 8월 30일 개정일자 : 2015년 9월 7일 접수일 : 2015년 9월 9일

제출 유형 : 일반 트랙

교신저자 : 신경식

1. 소개

파산 예측은 회계 및 금융 분야에서 광범위하게 연구되었습니다. 리스크 관리 측면에서 대출 결정 및 금융 기관의 수익성에 중요한 영향을 미칠 수 있습니다. 많은 연구자들은 보다 강력한 파산 예측 모델을 구축하는 데 집중해 왔습니다. 초기 연구에서는 파산 예측을 위해 다중 판별 분석(MDA)(Altman, 1968), 회귀 분석(Meyer and Pifer, 1970) 및 로짓 분석(Hamer, 1983; Ohlson, 1980)과 같은 통계 기법을 주로 사용했습니다.

그러나 최근 많은 연구에서 인공 신경망과 같은 인공 지능(AI) 접근 방식이 (ANN), 사례 기반 추론(CBR) 및 지원 벡터 머신(SVM)이 더 나은 성과를 보였습니다.

1990년대 이후로 MDA 및 로짓 분석과 같은 통계 기법이 비즈니스 분류 문제에 사용되었습니다.

센, 1997; 신과 한, 1999; 신과 한, 2001년; Shin et al., 2005).

많은 연구자들은 파산 예측 연구에서 ANN이 복잡하고 비선형적인 패턴 분류 문제를 다루는 데 통계적 방법보다 우수하다고 보고했습니다(Atiya, 2001; Boritz and Kennedy, 1995; Fletcher

및 Goss, 1993; 조 등, 1997; 레슈노와 스펙터, 1996; 오돔과 샤르다, 1990; 탐과 키앙, 1992; 윌슨과 샤다, 1994; Zhang et al, 1999). 특히 BP(back-propagation) 알고리즘으로 훈련된 MLP(Multilayer Perceptron) 네트워크는 방대한 연산 능력과 일반화 능력으로 인해 주로 사용된다.

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

많은 연구자들이 파산 예측 모델 개발에 중점을 두었지만 파산 유형을 분류하는 연구는 거의 없습니다. 이전의 파산 예측 모델은 일반적으로 기업이 파산할지 여부를 예측하는 데 관심이 있었습니다. 본 연구에서는 회계 및 재무 분야에 종사하는 연구자 및 실무자에게 기업 파산 유형에 대한 보다 유용한 정보를 제공하는 모델을 제안합니다. 이 모델은 수익성, 안정성 및 활동 지수 측면에서 특정 유형의 파산과 파산 발생을 예측합니다. 따라서 기업은 이를 사전에 예방할 수 있다.

파산 유형 예측을 위해 BPN과 SOM을 통합하는 모델을 설계했습니다.
한국 중소기업의 사례를 바탕으로
중견 건설사. 두 개의 ANN을 사용하는 하이브리드 방식의 프로세스는 다음 단계로 구성됩니다. 첫 번째 단계에서는 파산 예측을 위한 역전파 신경망(back-propagation neural network, BPN) 모델을 개발하고 파산 데이터를 여러 유형으로 나누는 SOM(Self-Organizing Map) 모델을 구성합니다. 이 두 단계는 순차적으로 수행되며, BPN 모델은 SOM 모델에서 예측한 파산 데이터를 사용하여 특정 파산 유형을 식별할 수 있습니다.

이 문서의 나머지 부분은 다음과 같이 구성됩니다. 2장에서는 본 논문의 연구 주제와 관련된 선행 연구 검토를 포함하여 파산 예측 및 파산 유형에 대한 이전 연구를 간략하게 설명합니다. 섹션 3은 다음을 제공합니다.

ANN의 방법론에 대한 설명
BPN 및 SOM이 포함됩니다. 섹션 4 설명
연구 데이터, 변수 및 실험 설계를 포함한 모델 구축 프로세스. 5장에서는 실험 결과 및 분석에 대해 설명합니다.
마지막으로 6장에서는 결론과 향후 연구 주제에 대해 논의한다.

2. 관련 업무

2.1 파산 예측

파산예측은 1960년대 후반부터 광범위하게 연구되어 왔다. 이 섹션에서는 전통적인 통계 기법과 파산 예측을 위한 ANN에 초점을 맞춰 이전 연구를 검토합니다. 대부분의 파산 예측 모델은 일반적으로 추정을 위해 재무 비율을 사용합니다. 기여

비버 (1966)
일변량 테스트를 사용하여 파산 예측에 크게 영향을 미칩니다. Altman(1968)의 MDA 기반 모형은 Z 점수에서 구한 값을 취하여 1년 전에 파산을 예측하는 높은 분류율을 보인다. MDA는 파산 예측 연구에 널리 사용되는 분석 기법입니다.

반면에 로짓 분석을 채택하면 MDA와 관련하여 논의된 모든 문제를 근본적으로 피할 수 있습니다. Ohlson(1980)은 파산 예측에 로짓 분석이 MDA보다 더 관련이 있음을 보여주었습니다. 이러한 통계적 방법은 선형성, 정규성 및

입력 간의 독립성. 따라서 1990년대 이후 많은 연구에서 ANN과 같은 데이터 마이닝 기술이 MDA 및 로짓 분석과 같은 통계적 방법보다 우수하다는 것이 입증되었습니다.

파산 예측을 위한 ANN의 첫 번째 채택은 Odom과 Sharda(1990)에 의해 연구되었습니다. 모델은 Altman(1968) 연구에서 사용된 5개의 입력 변수를 선택했습니다. 그들은 ANN 모델의 성능을 다음과 비교했습니다.

MDA의 것. 결과는 ANN

검증 세트에서 81.81%를 정확하게 분류한 반면 MDA는 74.28%만 얻었습니다.

Tam과 Kiang(1992), Leshno와 Spector(1996), Jo et al. (1997) 파산 예측을 위한 ANN과 판별 분석 간의 분류 정확도를 비교합니다. Tam and Kiang(1992)은 ANN 모델의 성능을 선형 판별 분석(LDA), 로짓 분석, 결정 트리 및 k-최근접 이웃의 성능과 비교했습니다.

그들은 1985년부터 1987년까지 수집된 은행의 파산 데이터를 사용했습니다. 경험적 결과는 신경망 모델이 훈련 데이터 세트 이전 1년 동안 다른 기술보다 우수한 것으로 나타났습니다. Leshno와 Spector(1996)는 다양한 데이터 범위, 학습 기술 및 파산 예측을 위한 반복 횟수 측면에서 다양한 ANN 모델의 예측 기능을 평가했습니다. ANN 모델의 성능도 판별분석 결과와 비교하였다.

그들은 ANN의 성능이 모델보다 더 정확하다.

기존 판별 모델. Joet al. (1997) 판별분석을 통계적 방법으로 적용

그리고 CBR과 ANN은 한국 기업의 파산 예측을 위한 인공 지능 방법입니다.

그들은 독립변수의 개수, 입력변수의 선택방법,

표준화. 결과는 ANN

다른 두 가지 방법을 능가했습니다.

ANN의 성능 비교와 파산 예측을 위한 로짓 분석에 대한 연구는 Fletcher and Goss(1993), Zhang et al. (1999), Tseng and Hu (2010).

Fletcher와 Goss(1993)는 ANN 모델을 로짓 모델과 비교했습니다. 그들의 모델은 표본 크기가 작기 때문에 세 가지 재무 비율을 사용했습니다. 그들은 18겹 교차 검증 방법을 사용하여

제안된 모델의 견고성을 평가합니다. 결과는 BPN 모델이 로짓 모델보다 더 정확한 분류를 제공함을 나타냅니다. Zhang et al. (1999) 또한 성능 측면에서 ANN을 로짓 모델과 비교하고 다음을 사용했습니다.

평가를 위한 5중 교차 검증 방법

신경 분류기의 견고성 그들은 ANN이 파산 예측을 위한 로짓 분석보다 분명히 우수하다는 결론을 내렸고 ANN 연구원들에게 많은 수의 무작위 시작 시드 세트를 사용하고 숨겨진 노드에서 실험할 것을 조언했습니다. Tseng and Hu(2010)는 4가지 다른 방법을 비교했습니다. 그들은 로짓 분석, 2차 간격 로짓 모델, BPN 및 방사형 기저 함수 네트워크(RBFN)를 사용하여 영국 기업의 파산을 예측했습니다. 그들은 RBFN이 분류 정확도 측면에서 다른 방법보다 우수하다고 결론지었습니다.

적용을 시도한 몇 가지 연구

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

파산 예측을 위한 감독되지 않은 신경망. Lee et al. (2005) 한국 기업의 파산 예측을 위한 분류 정확도 측면에서 감독된 신경망과 감독되지 않은 신경망을 비교했습니다. 판별 분석 및 로짓 분석도 벤치마크에 사용되었습니다. 경험적 결과는 출력변수가 주어졌을 때 BPN이 예측력을 가진다는 것을 보여주었다.

Boyacioglu et al.(2009)은 4가지 다른 신경망 모델, 지원 벡터 머신을 비교했습니다.

및 세 가지 다변량 통계 방법

터키 은행의 파산 예측. 그들은 네 가지 신경망 모델, 즉 MLP, SOM, 경쟁 학습 및 학습 벡터 양자화(LVQ)를 구성했습니다. 그들은 MLP 및 LVQ 신경망이

다른 방법.

각 기법의 장점을 접목하여 보다 강력한 기법을 개발하기 위해 ANN과 퍼지 및 유전 알고리즘(GA)과 같은 기타 기법을 이용한 통합 모델 개발을 진행하고 있습니다. Kim and Han(2001)은 파산 예측을 위한 퍼지 신경망 모델을 제안했습니다. 제안된 모델은 데이터 전처리를 위해 퍼지 멤버십 함수를 사용하였다.

결과는 퍼지의 성능을 보여주었다.

신경망은 신경망보다 정확합니다.

기존의 신경망. Hong and Shin(2003)은 파산 예측을 위한 최적 또는 거의 최적에 가까운 입력 변수를 선택하기 위해 GA를 채택했습니다.

그 결과 제안된 GA 기반 입력 선택 방법의 성능이 일변량 검정 및 단계적 방법과 같은 통계적 방법을 사용하는 모델보다 우수한 것으로 나타났습니다.

방법. Kim(2004)은 GA를 사용한 인스턴스 선택 접근 방식을 제안했습니다. GA는 레이어와 인스턴스 간의 최적 연결가중치를 찾는 데 사용됩니다.

결과는 제안한 방법이 분류 정확도를 높이는 데 유망한 것으로 나타났습니다.

최근 Lee and Choi(2013)는 다양한 산업의 한국 기업에 대한 파산 예측 모델을 제안했다. 그들은 건설, 소매 및 제조 산업에 대한 파산 예측 모델을 개발하기 위해 서로 다른 독립 변수 그룹을 선택했습니다. 그들은 전체 샘플뿐만 아니라 산업 샘플을 사용하여 ANN 모델의 예측 정확도를 MDA의 예측 정확도와 비교했습니다. 결과는 산업 표본을 사용한 성능이 전체 표본의 성능보다 더 정확함을 보여주었다. ANN의 성능은 MDA의 성능을 증가했습니다.

2.2 파산 유형의 분류

많은 연구자들이 파산 예측 모델 개발에 중점을 두었지만 데이터 마이닝 기술을 사용하여 파산 유형을 분류한 연구는 거의 없습니다. 대부분의 연구는 이전 문헌을 검토하거나 사례 연구를 수행하는 데 중점을 두었습니다. Argenti(1976)는 문헌 검토와 전문가 인터뷰를 통해 파산을 과정으로 조사했습니다. 재무적 요인과 비재무적 요인을 고려하여 창업기업, 신생기업, 성숙기업의 3가지 유형의 파산을 제안하였다.

파산의 첫 번째 유형은 경영이 제대로 이루어지지 않은 신생 기업을 의미합니다.

Nam-ok Jo ·Hyun-jung Kim ·Kyung-shik Shin

기술이나 성격의 조건. 주로 소규모 회사에서 발생하며 이러한 유형의 대부분의 회사는 파산합니다. 두 번째 유형은 젊은 기업에 적용되며 이 유형의 대부분의 기업도 훨씬 더 가파른 성장과 쇠퇴 후에 파산합니다.

이러한 기업은 성장 단계에서 운영 및 재무 관리가 적절하게 고려되지 않을 때 파산합니다. 마지막 유형의 기업은 환경 변화에 적절하게 대응하지 못하기 때문에 파산한다.

Moulton et al. (1996)은 기업 및 산업 성장 패턴에 따라 4가지 독특한 파산 유형을 제안했습니다. 실패의 첫 번째 유형

회사는 시장의 악화의 결과입니다.

이러한 유형의 기업은 쇠퇴하는 산업에서 매출이 감소하고 있습니다. 두 번째 유형은 시장의 부적절함으로 인해 발생합니다. 이러한 유형의 기업은 성장하는 산업에서 매출이 감소하고 있습니다. 파산 전 5년 동안 부채는 크게 늘어나고 자산은 감소합니다. 지난해 자산수익률 하락폭이 가장 컸다. 세 번째 유형은 시장 점유율을 위한 싸움에 참여합니다.

이들 기업은 쇠퇴하는 산업에서 매출이 증가하고 있습니다. 그러나 자산과 부채는 매출보다 훨씬 빠르게 증가합니다. 마지막 유형은 통제력 상실과 관련이 있습니다. 이러한 유형의 기업은 성장하는 산업에서 매출이 증가하고 있습니다. 매출, 자산, 부채 증가율이 어느 그룹보다 높고, 감소 속도도 빠릅니다.

이러한 유형을 통제 상실이라고 하지만 많은 파산 사례는 불행하고 특정한 사건의 결과입니다.

Ooghe and De Prijcker(2008)는 기업의 특성 간의 관계를 연구했으며,

파산의 원인과 파산 유형에 대한 통찰력을 얻기 위한 재정적 영향. 그들은 문헌 검토 및 사례 연구를 수행하고 기업의 성숙도 및 경영 특성 측면에서 파산 절차의 네 가지 유형(실패한 창업 기업, 야심찬 성장 기업, 눈부신 성장 기업, 무관심)을 제안했습니다.

설립된 회사.

파산 회사의 첫 번째 유형은 창업 회사입니다. 이들 기업은 경영에 관련 산업에서 요구되는 경영 기술과 경험이 심각하게 부족하기 때문에 파산하게 된다. 따라서 그들은 직면한 재정적 문제를 해결할 능력이 없습니다. 두 번째 유형은 야심찬 성장 기업으로, 위험을 추구하는 경영자에 의해 발생합니다. 이러한 기업은 가장 비용 효율적인 복귀 계획을 구현하는 데 필요한 재정적 수단이 부족합니다. 이들 기업 중 상당수는 환경적 요인에 적절히 대응하지 못해 파산한다.

세 번째 유형은 회사의 운영 구조를 간과하는 경향이 있는 눈부신 성장 회사입니다. 이로 인해 비용이 크게 증가하고 수익성이 악화됩니다. 기업은 가파른 성장과 쇠퇴 후 파산에 들어갑니다. 마지막 유형은 무관심한 기존 회사입니다. 이러한 기업은 수익성 있는 조직을 관리할 수 있는 능력이 있지만 헌신과 동기가 부족합니다. 그들은 그들의 태도 때문에 그들의 매출이 점차적으로 감소하고 있다는 것을 알아차리지 못한다.

회사에 대한 무관심. 이들 기업은 할 수 없다.

이해 관계자가 불신하기 때문에 생존

모든 재정적 수단이 소진되었습니다.

3. 방법론

ANN은 뉴런이라고 하는 조밀하게 상호 연결된 정보 처리 장치 세트로 구성됩니다.

이것은 복잡하고 비선형적인 패턴을 다루는 수많은 연구에서 광범위하게 수용되었습니다.

ANN은 지도 또는 비지도의 두 가지 학습 유형으로 분류할 수 있습니다. 지도 학습에는 입력과 해당하는 원하는 출력으로 구성된 훈련 세트가 있습니다. BPN은 지도 학습 유형에서 가장 널리 사용되는 모델입니다. 지도 학습과 달리 비지도 학습은 학습을 위한 입력으로만 구성됩니다. SOM은 널리 사용되는 신경망 기술입니다. 네트워크에는 입력 세트만 제공되며 원하는 출력은 제공되지 않습니다.

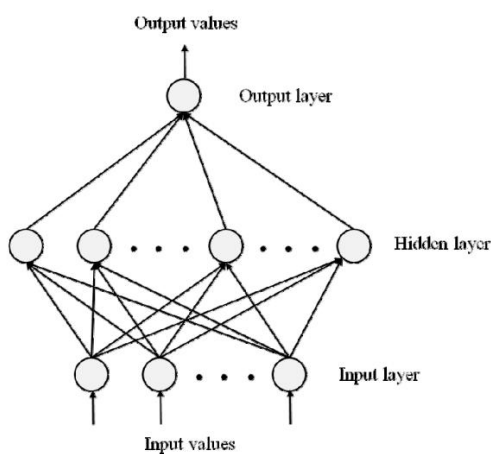
(Hecht-Nielsen, 1989)은 가장 널리 사용되는 신경망 기술 중 하나입니다. 일반적으로 입력 레이어, 출력 레이어 및 MLP라고 하는 하나 이상의 은닉 레이어와 같은 뉴런이라고 하는 처리 요소의 여러 레이어로 구성됩니다.

입력 계층은 네트워크에 공급되는 외부 정보를 나타냅니다. 출력 레이어(출력 맵이라고도 함)는 모델의 최종 솔루션을 생성합니다. 네트워크에는 복잡하고 비선형적인 패턴을 인식하는 하나 이상의 은닉층이 있습니다.

네트워크 내에서 각 뉴런은 연결된 숫자 가중치로 연결됩니다. BPN은 이러한 가중치를 조정하는 반복 프로세스를 통해 학습합니다. <그림 1>은 이진 분류 문제에 사용되는 3계층 MLP의 예를 보여준다. 뉴런은 (1)과 같이 입력 값의 가중치 합을 계산합니다.

3.1 역전파 신경망(BPN)

BPN(역전파 신경망) 네트워크



<그림 1> MLP의 아키텍처(Rumelhart et al., 1985)

$$\text{나쁜 } X = \sum_{i=1}^n x_i \quad (1)$$

여기서 X 는 뉴런에 대한 순 가중 입력, x_i 는 입력 값, w_i 는 입력 i 의 연결 가중치, n 은 뉴런 입력 수입니다. 가중치 합은 적절한 전달 함수에 의해 뉴런의 활성화 값으로 변환됩니다. 네 가지 일반적인 전달 함수가 있습니다: 단계 함수, 부호 함수, 시그모이드 함수 및 선형 함수. BPN에서 일반적으로 사용되는 전달함수는 S자 형태의 시그모이드 함수입니다. 시그모이드 함수는 다음 범위의 입력 값을 변환합니다.

$+\infty \sim -\infty$ 범위에서 합리적인 값으로

0과 1 사이. 이 기능을 가진 뉴런은

BPNN에서 사용됨(Negnevitsky, 2004). 전달 함수는 (2)에 나와 있습니다.

$$\text{그리고 } f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

BPNN에서 훈련의 목적은 반복적으로 뉴런 간의 가중치를 조정하는 것입니다.

오차를 최소화하는 방향. 오류는

평균 제곱 오류(MSE)와 같은 오류 측정을 사용하여 각 예측된 출력과 해당하는 원하는 출력 간의 차이를 비교하여 계산됩니다. 그런 다음 오류가 뒤로 전파됩니다. MSE는 (3)과 같이 정의할 수 있다.

$$\text{MSE} = \frac{1}{p} \sum_{p=1}^p (y_d - y)^2 \quad (3)$$

여기서 y_d 는 원하는 출력, y 는 실제 출력 값, p 는 훈련 패턴의 수입니다. 오차가 양수이면 값이 증가하는 방향으로 y 가 조정되고 오차가 음수이면 y 가 감소하는 경향이 있습니다. 분류 또는 예측 작업은 (4)와 같이 가중치 업데이트를 위해 다음 알고리즘을 사용하여 수행됩니다.

$$w(p+1) = w(p) + \eta \sum_{p=1}^p (y_d - y) \cdot x \quad (4)$$

학습률은 어디에 있습니까? 신경망의 훈련 과정은 수렴될 때까지 지정된 p 만큼 반복됩니다.

3.2 자기 조직화 지도(SOM)

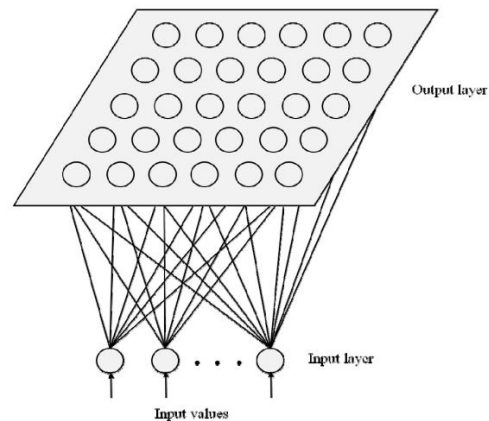
Kohonen(1982)이 개발한 자기 조직화 지도(SOM)는 고차원 데이터 세트의 패턴을 탐색하고 시각화하기 위한 모델입니다.

SOM은 피드포워드 ANN입니다. 다르다

SOM은 입력 값만 사용하고 원하는 출력이 제공되지 않는다는 사실에 의해 지도 학습에서 파생됩니다. <그림 2>와 같이 SOM은 Input Layer와 Output Layer의 두 가지 Layer로 구성되어 있다. 입력 층에는 변수 수만큼 많은 뉴런이 있으며 그 기능은 정보를 반영하는 것입니다. 출력 레이어는 2차원 배열 뉴런을 나타냅니다. 입력 계층의 각 뉴런은 출력 계층의 각 뉴런과 완전히 상호 연결됩니다. 각 뉴런 i 에는 연관된 d 차원 가중치 벡터 w_i 가 있습니다.

$w_i = [w_{i1}, w_{i2}, \dots, w_{id}]$, 출력 레이어. 이것 벡터를 참조 벡터라고 합니다.

최종 클러스터를 해석하는 데 사용됩니다.



<그림 2> SOM의 아키텍처

차원 d 는 입력과 동일합니다.

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

벡터 차원입니다. SOM은 모든 입력 벡터가 처리될 때까지 반복적으로 학습됩니다. 각 훈련 단계에서 샘플 벡터 x 는 입력 데이터 세트에서 무작위로 선택됩니다. x 와 각 가중치 벡터 w_i 사이의 거리는 SOM에서 널리 사용되는 유클리드 거리를 사용하여 계산됩니다. 연결 가중치가 x 와 더 높은 유사도를 갖는 단위를 BMU(Best-Matching Unit)라고 하며 c 로 표시됩니다 (Chang and Liao, 2006; Moreno et al., 2006).

$$\|x - w_c\| \leq \|x - w_i\| \quad (5)$$

BMU를 찾은 후 SOM의 연결 가중치가 조정됩니다. BMU와 토폴로지 이웃은 입력 벡터에 더 가깝게 이동됩니다(Alhoniemi et al., 1999). 뉴런 i 의 가중치 벡터에 대한 SOM 업데이트 규칙은 (6)과 같습니다.

$$w_i(t+1) = w_i(t) + \eta(t) h_i(t) [x(t) - w_i(t)] \quad (6)$$

여기서 t 는 시간이고, $\eta(t)$ 는 시간의 감소 함수인 학습률이고, $h_i(t)$ 는 c 주변의 이웃 커널입니다.

$$h_i(t) = \exp\left(-\frac{\|w_i - w_c\|^2}{2\sigma^2}\right) \quad (7)$$

r_c 와 r_i 는 SOM 그리드에서 뉴런 c 와 i 의 위치입니다. $\eta(t)$ 와 $h_i(t)$ 는 시간이 지남에 따라 단조롭게 감소합니다.

SOM은 분류와 같은 다양한 응용 분야에서 광범위하게 사용되었습니다.

(Corridoni et al., 1996; Deschenes and Noonan, 1995; Li et al., 2011; Moreno et al., 2006; Silver and Shmoish, 2008), 클러스터링(Mangiameli et al., 1996; Murtagh, 1995; -Brío 및 Serrano-Cinca, 1993) 및 예측(Van Der Voort et al., 1996). SOM은 보다 전통적인 통계 기법의 선형성 또는 정규성에 대한 제한적인 가정 없이 데이터 세트에서 클러스터를 식별하는 이점이 있습니다(Moreno et al., 2006; Mostafa, 2009).

4. 모델 개발

4.1 데이터와 변수

샘플 데이터는 106개의 재무 재무제표에서 파생된 비율과

해당 한국 중소 건설 회사. 사용 가능한 총 표본 수는 2002년부터 2007년까지 5,000개 기업(도산 사례 2,500건 및 비도산 사례 2,500건)을 포함합니다. 우리는 모든 용매 기업 중에서 2,500개의 비도산 기업을 무작위로 선택합니다. 데이터 세트는 훈련 세트, 테스트 세트 및 모든 데이터의 각각 60%, 20% 및 20%의 검증 세트의 세 가지 하위 세트로 분할됩니다. 훈련 세트는 다음을 위해 사용됩니다.

모델 구성 및 테스트 세트를 사용하여

학습을 위한 정지 조건. 검증 데이터는 모델의 유효성을 테스트하는 데 사용됩니다.

입력 변수 선택 프로세스의 두 단계를 수행합니다. 첫 번째 단계에서는 독립표본 t-검정과 MDA 단계적 방법과 같은 통계적 검정을 사용하여 변수를 선택합니다. 이것

방법은 일변량 검정을 만족하는 변수를 선택한 다음 MDA 단계적 방법을 사용하여 차원을 줄이는 중요한 변수를 선택합니다.

2단계에서는 1단계에서 선정된 변수가 수익성, 안정성, 활동성 등 다양한 지표군에 고르게 분포되어 있는지 검토하고 의미와 중복되는 변수를 제거하여 8개의 입력변수를 선정한다. 본 연구를 위해 선정된 변수는 <표 1>과 같다.

<표 1> 변수의 정의

지수 변수	정의
수익성	X1 총자산 대비 이익잉여금
	판매에 X2 EBITDA
안정	X3 빠른 비율
	X4 총 자산에 대한 주주 자본
	X5 판매에 대한 금융 비용
	X6 총 자산 대비 총 차입금
활동	X7 총 자산 회전율
	X8 고정 자산 회전율

4.2 실험 설계

본 연구에서는 파산 유형 예측을 위해 두 개의 ANN을 사용하여 모델을 설계합니다. 첫 번째 단계에서는 파산 예측을 위한 BPN 모델을 개발하고 파산 데이터를 여러 유형으로 나누는 SOM 모델을 구성합니다. 두 번째 단계에서는 BPN 모델을 검증 세트에서 활용하지 않는 검증 세트에 적용하여 기업의 파산을 예측합니다.

모델의 개발. 이것은 허용합니다

BPN 모델에 의해 예측된 파산 데이터를 사용하여 특정 유형의 파산을 식별합니다. 본 연구에서 사용된 재무비율은 BPN과 SOM의 학습능력을 향상시키고 결과 클러스터의 특성을 쉽게 해석하기 위해 평균과 표준편차를 이용하여 표준화하였다.

4.2.1 BPN을 통한 파산 예측

파산 예측 모델을 구축하려면 네트워크 아키텍처와 학습 매개변수를 정의해야 합니다. 즉, 먼저 최적의 네트워크를 찾는 다음 사용할 학습 알고리즘을 결정해야 합니다. 마지막으로 네트워크의 가중치는 훈련 예제 세트에서 초기화되고 업데이트됩니다. 모델 설계에 대한 특정 원칙이 없기 때문에 최적의 네트워크 아키텍처는 일반적으로 실험을 통해 결정됩니다.

BP 알고리즘으로 훈련된 MLP 네트워크를 선택합니다. 시그모이드 전달 함수는 은닉 노드와 출력 노드에서 사용됩니다. 숨겨진 노드의 수는 실험에 의해 결정됩니다. 입력 노드의 수는 입력 변수의 수와 동일합니다. 하나의 출력 노드는 이진 분류에 사용됩니다. 파산은 출력으로 정의되며 출력의 범위는 [0, 1]입니다.

학습률은 0.3으로 설정되고 운동량은 다음과 같습니다. 0.5로 설정합니다.

4.2.2 파산 유형 분류 숨

재무 비율은 입력 변수로 사용됩니다.

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

파산 유형을 분류하는 지도를 생성합니다.
입력 계층에 8개의 노드가 있는 자체 구성 네트워크가 사용됩니다.
2차원 출력 맵의 크기 선택은 일반적으로 명확한 규칙이 없기 때문
에 실험에 의해 선택됩니다.

따라서 그리드에서 다양한 너비와 길이를 설정하여 실험한 다음 최상의 SOM 클러스터링 구성을 결정합니다.

SOM 교육은 두 단계로 나뉩니다. 첫 번째 단계는 대략적인 추정 단계로 데이터에서 큰 패턴을 찾습니다. 두 번째 단계는 조정 단계로 데이터의 정제된 기능을 모델링하기 위해 맵을 조정합니다. 각 단계마다 학습 매개변수, 이웃, 초기 학습률, 주기 설정이 필요합니다. 이웃의 시작 크기는 훈련 중에 승리한 유닛에 의해 업데이트되는 근거리 유닛의 수를 찾기 위해 설정됩니다(IBM, 2012). 첫 번째 단계에서는 이웃 수를 5로 설정하고 초기 학습률을 0.5로 설정합니다. 다음 단계에서는 이웃의 수를 3으로 설정하고 초기 학습률을 0.2로 설정합니다. 이웃 크기는 5 이웃에서 시작하여 첫 번째 단계에서 4로 감소하고 이웃 크기는 3에서 시작하여 1로 감소합니다. 학습률은 첫 번째 단계에서 0.5에서 시작하여 0.5에서 시작하여 0으로 감소합니다. 따라서 첫 번째 단계의 이웃 수와 학습률은 두 번째 단계의 이웃 수와 학습률보다 커야 합니다. 학습률은 가중치 요인으로 시간이 지남에 따라 훈련 중에 선형적으로 감소합니다. 주기의 수는 각 단계에 대해 할당된 전송 데이터 수에 대해 계속됩니다.

훈련. 사이클 수는 첫 번째 단계에서 80, 두 번째 단계에서 200으로 할당됩니다.

5. 결과 및 분석

파산 유형 예측을 위해 두 개의 ANN을 통합하는 모델을 개발했습니다. 첫 번째는 파산 예측을 위한 BPN 모델이었습니다.

두 번째는 분할하는 SOM 모델이었습니다.

파산 데이터를 여러 유형으로 나눕니다. BPN 모델 16개의 히든 노드로 최고의 성능을 보여주었습니다.
분류 정확도. 결론적으로 BPN의 파산 예측 모델은 훈련 세트의 경우 78.06%, 테스트 세트의 경우 77.30%로 올바르게 분류되었습니다. 그런 다음 실제 파산 데이터 세트를 사용하여 파산 유형을 식별하는 SOM 모델을 구성했습니다. SOM 학습을 위해 5 x 3 정사각형 그리드가 고려되었습니다. 그 결과 SOM 클러스터링 결과에 따라 8개의 재무비율을 이용하여 5개의 클러스터를 분류하였다. <표 2>는 주파수와

각 클러스터의 비율입니다.

<표 2> 모델에 대한 결과 클러스터의 크기
건설 세트

클러스터	빈도	%
1	555	27.75%
2	290	14.50%
3	290	14.50%
4	328	16.40%
5	537	26.85%
총	2000년	100.0%

입력변수 선정과정을 통해 선정된 입력변수들의 평균을 구함

<표 3> 모델 구성 세트의 입력 변수에 대한 각 클러스터의 평균

변수	평균					
	유형 1	유형 2	유형 3	유형 4	유형 5	총
X1	-0.645	-0.740	-0.267	-0.474	0.084	-0.380
X2	0.114	-0.333	-0.067	-0.556	-0.265	-0.189
X3	-0.487	-0.631	-0.292	-0.568	-0.096	-0.388
X4	-0.869	-1.269	-0.072	-0.814	0.491	-0.437
X5	1.722	0.257	0.447	-0.469	-0.475	0.376
X6	1.256	0.985	0.432	0.039	-0.439	0.442
X7	-0.834	0.077	-0.456	0.631	0.160	-0.140
X8	-0.608	0.192	-0.347	0.712	0.170	-0.029

파생된 클러스터의 특성을 해석하기 위해 각 클러스터에 대해, <표 3>은 모든 입력변수에 대한 각 군집의 평균을 나타낸다.

각 군집은 파산기업의 자료를 통해 분류된 파산유형을 나타내며, 입력변수는 각 군집의 의미를 해석하는데 있어서 재무비율을 나타낸다. 유형 1의 재무 비율은 클러스터링 결과에 따라 매출 대비 EBITDA(이자, 세금, 감가상각 전 이익)를 제외하고 재무제표가 열 등합니다. 첫째, 이 유형은 판매에 대한 재정적 비용이 높고 총 자산에 대한 총 차입금이 있습니다. 이것은 기업이 장기적으로 운영되는 안정적인 자산을 가지고 있지 않다는 것을 의미합니다. 이러한 기업은 가장 비용 효율적인 복구 계획을 수행하는 데 필요한 자산이 부족합니다.

둘째, 자산을 사용하여 이익을 축적할 수 있는 기업의 능력을 측정하는 총자산 대비 이익잉여금이 낮고 총자산 대비 자기자본이 낮습니다. 마지막으로, 이 유형은 낮은

총 자산 회전율 및 고정 자산 회전율. 이것들 비율은 활동 지수를 나타내며 기업이 총 자산을 얼마나 잘 사용하고 있는지 나타냅니다.

매출을 창출하기 위한 고정 자산. 낮은 비율은 기업이 상대적으로 적은 총자산이나 고정자산으로 매출을 효과적으로 창출하지 못하고 있음을 나타냅니다. 따라서 기업에서는 불가능하다.

정상 복구를 위해 이 클러스터에 포함 부채 능력의 손실과 최고 수준의 신용 위험으로 인해 회사.

유형 2는 당좌비율이 낮고 총자산 대비 자기자본이 낮고 총 자산 대비 총차입금이 높다. 첫째, 당좌비율이 낮다는 것은 기업이 자산을 사용하여 유동부채를 신속히 처분할 능력이 없다는 것을 나타냅니다.

둘째, 총자산 대비 낮은 자기자본 기업의 재무구조를 나타냅니다. 안좋다. 마지막으로 위에서 언급한 바와 같이 총자산 대비 총차입금이 높다는 것은 기업이 장기적으로 운영되는 안전한 자산이 없다는 것을 의미합니다. 따라서 이 클러스터에 속한 기업이 정상적인 기업으로 회복하는 것은 거의 불가능합니다.

유형 3은 총 자산이 약간 낮습니다. 회전율 및 고정 자산 회전율. 이러한 비율 비즈니스가 매출을 창출하기 위해 총 자산 또는 고정 자산을 얼마나 잘 사용하고 있는지 나타냅니다. 처럼

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

<표 4> 파산 유형의 특성

파산 유형	설명
유형 1: 심각한 파산	다른 클러스터에 비해 열등한 재무제표 총자산 대비 이익잉여금이 낮고 총자산 대비 자기자본이 낮음 판매에 대한 높은 금융 비용 및 총 자산에 대한 총 차입금 낮은 총 자산 회전을 및 고정 자산 회전을
유형 2: 안정성 부족	다른 자수에 비해 상대적으로 낮은 안정성 자수 총 자산 대비 높은 총 차입금 총자산 대비 자기자본이 낮고 당좌비율이 낮음
유형 3: 활동 부족	다른 자수에 비해 상대적으로 낮은 활동 자수 낮은 총 자산 회전을 및 고정 자산 회전을
유형 4: 수익성 부족	다른 자수에 비해 상대적으로 낮은 수익성 자수 총자산 대비 낮은 이익잉여금 및 매출 대비 EBITDA
유형 5: 회복 가능한 파산	다른 클러스터에 비해 상대적으로 우수한 재무제표 수익성을 나타내는 총 자산 대비 이익 잉여금 이 높습니다. 높은 당좌비율 및 총자산 대비 자기자본 판매에 대한 낮은 금융 비용 및 총 자산에 대한 총 차입금

유형 1에서 언급된 활동 자수가 낮으면 기업이 상대적으로 적은 총 자산으로 매출을 창출하는 데 효율적이지 않다는 것을 나타냅니다.

또는 고정 자산.

유형 4는 수익성 지표를 나타내는 총자산 대비 이익잉여금이 낮고 매출 대비 EBITDA가 낮습니다. 첫째, 총자산 대비 이익잉여금은 자산을 사용하여 이익을 축적할 수 있는 기업의 능력을 측정합니다. 이 비율은 회사의 나이를 반영하는 수익성을 나타냅니다. 신생 기업이나 신생 기업은 총 자산 대비 이익 잉여금이 낮습니다[1]. 둘째, 매출 대비 EBITDA는 매출을 비교하여 기업의 수익성을 측정합니다.

그러나 이러한 유형은 파산한 기업임에도 불구하고 안정적인 재무구조와 활동성 측면에서 장점이 있다.

유형 5는 파산하더라도 다른 유형의 파산에 비해 상대적으로 재무 상태가 양호한 기업을 포함합니다.

이 유형은 파산 기업의 매출 대비 EBITDA를 제외한 모든 재무 비율에 대해 양호한 재무제표를 가지고 있습니다. 이 유형의 기업은 총 자산 대비 이익 잉여금이 높습니다. 특히 총자산 대비 자기자본이 상대적으로 높고 매출에 대한 금융비용과 총자산에 대한 총차입금이 다른 클러스터에 비해 상대적으로 낮은 편이다. 그러나 재무구조가 비교적 양호한 편이라 할지라도 경제환경의 변화에 따라 재무상태가 악화될 수 있다.

각 파산 유형의 특성은 다음과 같습니다.

<표 4>에 요약되어 있다.

BPN 및 SOM 모델을 개발한 후 유효성을 테스트하기 위해 검증 세트에 모델을 적용했습니다. 그런 다음 검증 세트에 대한 기업의 파산을 예측하고 BPN 및 SOM 모델의 파산 예측 데이터를 기반으로 특정 파산 유형을 분류했습니다.

검증을 위해 76.30%로 정확하게 분류된 BPN 기반 파산 예측 모델 세트. 총 513개의 기업이 다음으로 분류되었습니다. 파산했고 487개 기업이 용제기업으로 분류됐다. 513개 기업의 파산 데이터를 사용하여 특정 파산 유형을 분류했습니다. 따라서 우리는 먼저 파산 회사를 예측한 다음 실제 세계에서 새로운 데이터가 제공될 때 특정 파산 유형을 알 수 있습니다. <표 5>는 주파수와 검증 세트에 대한 각 클러스터의 비율.

<표 5> 결과 클러스터의 크기 검증 세트

클러스터	빈도	%
1	78	15.21%
2	100	19.49%
3	72	13.84%
4	97	18.91%
5	167	32.55%
총	513	100.0%

또한 5개 유형의 파산 유형 간의 기술적인 차이를 보여주기 위해 평균을 계산했습니다.

모델에 사용되지 않는 검증 세트
건설. <표 6>은 8개 변수에 대한 유형별 평균을 나타낸다.

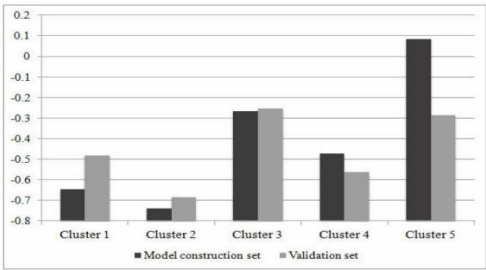
<그림 3>은
모델 구성 세트와
검증 세트. SOM에서 구축한 파산유형분류모형에 검증집합을 적용하여 유형별 평균재무비율은 이익잉여금을 제외하고 모형구성집합과 유사함을 알 수 있었다.

판매 및 EBITDA를 판매로 연결합니다.

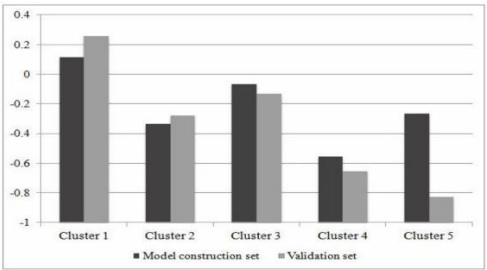
<표 6> validation set의 입력 변수에 대한 각 클러스터의 평균

변수	평균					
	유형 1	유형 2	유형 3	유형 4	유형 5	총
X1	-0.483	-0.684	-0.255	-0.562	-0.285	-0.475
X2	0.257	-0.277	-0.131	-0.656	-0.829	-0.241
X3	-0.450	-0.624	-0.305	-0.542	-0.126	-0.432
X4	-0.830	-1.233	0.355	-0.936	0.515	-0.602
X5	1.167	0.393	0.563	-0.510	-0.201	0.565
X6	1.202	0.865	0.418	0.117	-0.271	0.594
X7	-0.829	0.038	-0.457	0.595	-0.011	-0.212
X8	-0.585	0.111	-0.298	0.573	0.318	-0.052

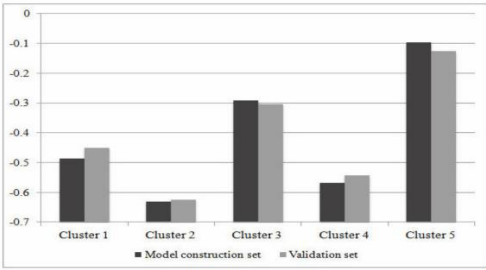
하이브리드 인공 신경망 모델을 사용한 파산 유형 예측



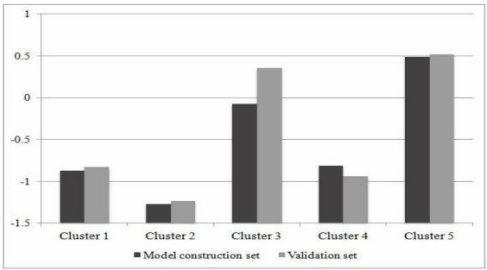
(a) Retained earnings to total assets



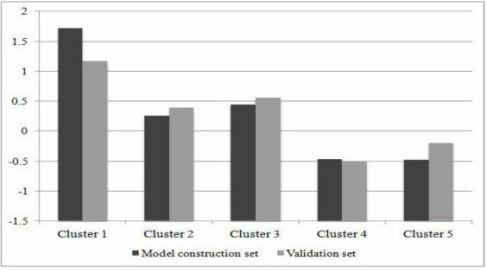
(b) EBITDA to sales



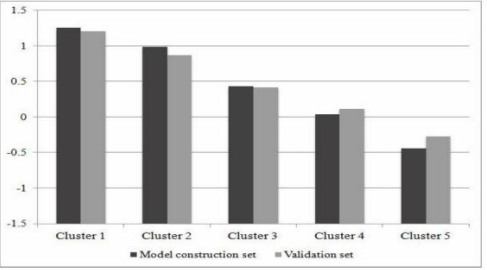
(c) Quick ratio



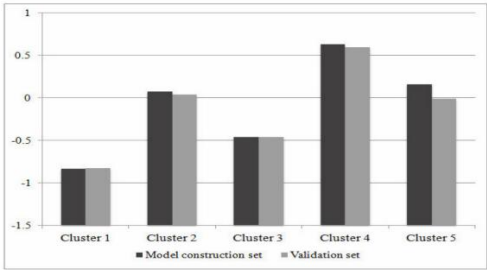
(d) Stockholder's equity to total assets



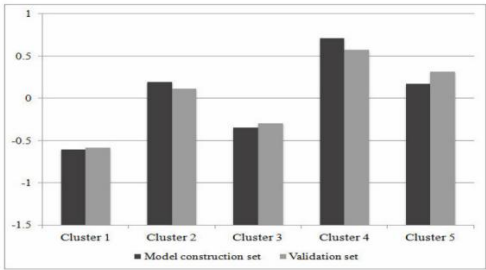
(e) Financial expenses to sales



(f) Total borrowings to total assets



(g) Total asset turnover



(h) Fixed asset turnover

<그림 3> A 모델 구성 세트와 A 모델의 평균 차이
검증 세트

6. 결론

파산의 예측은 회계 및 금융 분야에서 광범위하게 연구되었습니다. 그러나 데이터 마이닝 기법을 이용하여 특정 파산 유형을 분류한 연구는 거의 없었다. 파산 유형을 예측하기 위해 두 개의 ANN을 사용하는 하이브리드 접근 방식을 제안했습니다. 제안된 모델은 기업의 도산을 미리 예측할 수 있을 뿐만 아니라 특정 유형의 도산을 예측할 수 있습니다.

샘플 데이터는 재무제표 및 해당 한국의 무감사 건설 회사에서 파생되었습니다. 본 연구에서는 BPN을 사용하여 파산을 예측하는 모델을 구축하고 SOM을 사용하여 파산을 여러 유형으로 구분했습니다. 구축된 BPN 및 SOM 모델을 채택하여 실제 세계에서 새로운 데이터에 대한 파산 유형을 예측할 수 있습니다. 실증적 결과는 5가지 유형의 파산 유형이 8가지 재무비율에 따라 각기 다른 특성을 가짐을 보여주고 있다. 연구 결과를 바탕으로 관련 분야에 종사하는 연구자 및 실무자는 기업의 파산 유형에 대해 보다 유용한 정보를 얻을 수 있다.

이 연구에는 몇 가지 제한 사항이 있습니다. 먼저 SOM 클러스터링 결과는 매개변수 값에 큰 영향을 받기 때문에 시행착오를 통해 다른 조건을 설정하여 2차원 출력 맵의 크기, 이웃, 학습률, 주기 등 관련 매개변수를 정의해야 합니다. 작업 프로세스의 복잡성. 따라서 이 접근 방식은

에서 의미 있는 실험 결과를 도출

SOM 클러스터링은 파산에 따라 최상의 클러스터를 분석하는 데 시간과 비용이 많이 듭니다.

유형.

둘째, 기업의 재무비율을 고려하여 파산유형을 분류하였다.

파산을 정확하게 예측하고 파산의 종류를 의미 있게 나누는 입력변수를 선택하는 것이 중요하다. 추가 연구에서,

비재무적 요소도 포함해야 합니다.

기업의 규모, 업종, 나이 등. 따라서 질적 요인을 통합하고 전문가의 도메인 지식을 반영하여 현실적이고 효과적인 클러스터링 결과를 얻을 수 있습니다.

참고문헌

Alhoniemi, E., J. Holmén, O. Simula 및 J.

Vesanto, "자기 조직화 지도를 이용한 프로세스 모니터링 및 모델링", 통합 컴퓨터 지원 공학, Vol.6, No.1(1999), 3~14.

Atiya, AF, "신경망을 이용한 신용위험 파산 예측: 조사 및 새로운 결과," IEEE Transactions on Neural Networks, Vol.12, No.4(2001), 929~935.

알트만, EI, "재무비율, 판별분석과 기업도산예측," 금융저널, Vol.23, No.4(1968), 589~609.

Argenti, J., 기업 붕괴: 원인 및 증상, 런던: McGraw-Hill, 1976.

Beaver, WH, "실패의 예측 변수로서의 재무 비율", Journal of Accounting Research,

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

- 권. 4(1966), 71~111.
- Boritz, JE 및 DB Kennedy, "비즈니스 실패 예측을 위한 신경망 유형의 효율성", 응용 프로그램 전문가 시스템, Vol.9, No.4(1995), 503~512.
- Boyacioglu, MA, Y. Kara 및 Ö. K. Baykan, "신경망, 지원 벡터 머신 및 다변수 통계 방법을 사용한 은행 재정 실패 예측: 터키의 저축성 예금 보험 기금(SDIF) 아체 은행 샘플의 비교 분석," Expert Systems with Applications, Vol.36, 제2호(2009), 1685~1694.
- Chang, PC 및 TW Liao, "반도체 제조 이론에서 흐름 시간 예측을 위한 SOM 및 퍼지 규칙 기반 결합"
- 응용 소프트웨어 컴퓨팅, Vol.6, No.2(2006), 198~206.
- Corridoni, JM, A. Del Bimbo, L. Landi, "다객체 코호넨 네트워크를 이용한 3차원 객체 분류," Pattern Recognition, Vol.29, No.6(1996), 919~935.
- Deschenes, CJ 및 J. Noonan, "과도 분류를 위한 퍼지 코호넨 네트워크
특징 추출을 위한 웨이블릿 변환 사용", Information Sciences, Vol.87, No.4(1995), 247~266.
- Fletcher, D. and E. Goss, "신경망을 이용한 예측: 파산 데이터베이스를 이용한 응용," Information and Management, Vol.24, No.3(1993), 159~167.
- Hamer, M., "실패 예측: 대체 통계 방법 및 변수 집합에 대한 분류 정확도의 민감도," 회계 및 공공 정책 저널, Vol.2, No.4(1983), 159~167.
- Hecht-Nielsen, R., "역전파 이론
신경망," 신경망에 관한 IJCNN 국제 합동 회의 1989, (1989), 593~605.
- 홍성희. 및 K.-S. 신, "인공 신경망 모델링을 위한 GA 기반 입력 선택 방법 활용: 파산 예측에 응용," 지능정보시스템학회지, 제9권, 제1호(2003), 227~249.
- IBM, IBM SPSS Modeler 15 사용자 안내서, 2012.
- Rumelhart, DE, GE Hinton 및 RJ Williams, "오류 전파를 통한 내부 표현 학습", CALIFORNIA UNIV SAN DIEGO LA JOLLA INST FOR 인지 과학, 번호 ICS-8506(1985).
- 조현, 한일, 이혜, "사례기반추론, 신경망, 판별분석을 이용한 파산예측," Expert Systems with Applications, Vol.13, No.2(1997), 97~108 .
- 김경지 "인공신경망 파산예측을 위한 인스턴스 선택을 이용한 데이터마이닝," 지능정보시스템학회지, Vol.10, No.1(2004), 109~123.
- Kim, KJ, HanIG, "Fuzzy Neural Networks를 이용한 파산 예측," 지능정보시스템학회지, Vol.7, No.1(2001), 135~147.
- Kohonen, T. (1982). 위상학적으로 올바른 기능 맵의 자체 조직화. 생물학 사이버네틱스, Vol.43, No.1, 59~69.
- Lee, K., D. Booth, P. Alam, "한국 기업의 파산 예측에 있어 감독 신경망과 감독되지 않은 신경망의 비교," 응용 프로그램이 있는 전문가 시스템, Vol.29,

- No.1(2005), 1~16.
- Lee, S., WS Choi, “역전파 신경망과 다변수 판별 분석을 이용한 다산업 도산 예측 모델,” Expert Systems with Applications, Vol.40, No.8(2013), 2941~2946.
- Leshno, M. and Y. Spector, "신경망 예측 분석: 파산 사례", 신경 컴퓨팅, Vol.10, No.2(1996), 125~247.
- Li, DC, WL Dai 및 WT Tseng, “차별화를 구축하기 위해 고객 특성을 분석하는 2단계 클러스터링 방법
- 고객관리: 섬유제조업 사례”, Expert Systems with Applications, Vol.38, No.6(2011), 7186~7191.
- Maher, JJ, TK Sen, "신경망을 이용한 채권등급 예측: 로지스틱 회귀와의 비교," Intelligent Systems in Accounting, Finance and Management, Vol.6, No.1(1997), 59~72.
- Mangiameli, P., SK Chen, D. West, “SOM 신경망과 계층적 클러스터링 방법의 비교, European Journal of Operational Research, Vol.93, No.2(1996), 402~417.
- Martín-del-Brío, B. 및 C. Serrano-Cinca, "데이터 분석 및 표현을 위한 자기 조직화 신경망: 일부 재무 사례", 신경 컴퓨팅 및 응용, Vol.1, No.3(1993), 193~206.
- Meyer, PA 및 HW Pifer, "은행 부실 예측", Journal of Finance, Vol.25, No.4(1970), 853~868.
- Moreno, D., P. Marco 및 I. Olmeda, "자체 구성 맵은 스페인 뮤추얼 펀드의 분류를 개선할 수 있습니다."
- European Journal of Operation Research, Vol.174, No.2(2006), 1039~1054.
- Mostafa, MM, "Shade of green: 자기 조직화 지도를 사용한 쿠웨이트의 녹색 소비자의 심리학적 세분화," Expert Systems with Applications, Vol.36, No.8(2009), 11030~11038.
- Moulton, WN, H. Tomas 및 M. Pruett, "비즈니스 실패 경로: 환경 스트레스 및 조직 대응," 관리 저널, Vol.22, No.4(1996), 571~595.
- Murtah, F., "인접성 제약 클러스터링을 사용하여 Kohonen 자체 구성 기능 맵 해석," Pattern Recognition Letters, Vol. 16, No.4(1995), 399~408.
- Negnevitsky, M., 인공지능. (2판), 피어슨 교육, 2004.
- Odom, M. 및 R. Sharda, "파산 예측을 위한 신경망 모델", 신경망에 대한 IJCNN 국제 공동 회의 1990 회보, (1990), 163~168.
- Ohlson, J., “재무비율과 파산의 확률적 예측,” 회계학회지, Vol.18(1980), 109~131.
- Ooghe, H. and S. De Prijcker, “기업도산의 실패과정과 원인: 유형학,” Management Science, Vol.46(2008), 223~242.
- Shin, K.-s. TS Lee, H.-j. 김, “도산예측모형에서 서포트 벡터 머신의 응용,” Expert Systems with Applications, Vol.28, No.1(2005), 127~135.
- Shin, K.-s. and I. Han, “유전자알고리즘이 뒷받침하는 회사 채 평가의 사례기반 추론,” Expert Systems with applications, Vol.16, No.2(1999), 85~95.

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

- Shin, K.-s. and I. Han, "회사채 등급에 대한 귀납적 색인을 사용한 사례 기반 접근," Decision Support Systems, Vol.32, 제1호(2001), 41~52.
- Silver, H. and M. Shmoish, "비지도 클러스터링 모델을 사용한 정신분열병 환자 및 건강한 개인의 인지 수행 분석," 정신의학 연구, Vol.159, No.1(2008), 167~179.
- Tam, KY 및 MY Kiang, "신경망의 관리 응용: 은행 부실 예측 사례" Management Science, Vol.38, No.7(1992), 926~947.
- Tseng, FM 및 YC Hu, "4가지 파산 예측 모델 비교: Logit, 2차 간격 로짓, 신경 및 퍼지 신경망", 응용시스템 전문가, Vol.37, No.3(2010), 1846~1853.
- Van Der Voort, M., M. Dougherty 및 S. Watson, "교통 흐름을 예측하기 위해 아리마 시계열 모델과 코호넨 지도 결합", 교통 연구 Part C: Emerging Technologies, Vol.4, No.5(1996), 307~318.
- Wilson, R. and R. Sharda, "신경망을 이용한 파산 예측," Decision Support Systems, Vol. 11, No.5(1994), 545~557.
- Zhang, G., YM Hu, EB Patuwo 및 CD Indro, "도산예측에서의 인공신경망: 일반적인 통과 교차검증분석", European Journal of Operational Research, Vol.116, No.1(1999), 16~32.

국문요약

하이브리드 인공지능망 모델을 이용한 부도 유형 예측

1)조남옥* ·김현정* ·신경식**

부도 예측은 회계와 재무 분야에서 꾸준히 연구되고 있는 분야이다. 초기에는 주로 다중판별분석(multiple discriminant analysis)와 로짓 분석(logit analysis)과 같은 통계적 방법을 이용하였으나, 1990년대 이후에는 경영 분야의 분류 문제를 위해 많은 연구자들이 인공지능망(back-propagation neural network), 사례기반추론(case-based reasoning), 서포트 벡터 머신(support vector machine) 등과 같은 인공지능을 통한 접근법을 이용하여 통계적 방법보다 분류 성과 측면에서 우수함을 입증해왔다. 기존의 기업의 부도에 관한 연구에서 많은 연구자들이 재무비율을 이용하여 부도 예측 모델을 구축하는 것에 초점을 맞추어왔다. 부도예측에 관한 연구가 꾸준히 진행되고 있는 반면, 부도의 세부적인 유형을 예측하여 제시하는 것에 대한 연구는 미흡한 실정이었다. 따라서 본 연구에서는 수익성, 안정성, 활동성 지표를 중심으로 국내 비외감 건설업 기업들의 부도 여부뿐만 아니라 부도의 세부적인 유형까지 예측 가능한 모델을 개발하고자 한다. 본 연구에서는 부도 유형을 예측하기 위해 두 개의 인공지능망 모델을 결합한 하이브리드 접근법을 제안하였다. 첫 번째 인공지능망 모델은 부도예측을 위한 역전파 인공지능망을 이용한 모형이며, 두 번째 인공지능망 모델은 부도 데이터를 몇 개의 유형으로 분류하는 자기조직화지도(self-organizing map)를 이용한 모형이다. 실험 결과를 통해 정의된 5개의 부도 유형인 심각한 부도(severe bankruptcy), 안정성 부족(lack of stability), 활동성 부족(lack of activity), 수익성 부족(lack of profitability), 회생 가능한 부도(recoverable bankruptcy)는 재무 비율에 따라 유형별로 상이한 특성을 갖는 것을 확인할 수 있었다. 본 연구 결과를 통해 신용 평가 분야의 연구자와 실무자들이 기업의 부도의 유형에 대한 유용한 정보를 얻을 것으로 기대한다.

주제어 : 부도 예측, 부도 유형 분류, 인공지능망, 역전파 인공지능망, 자기조직화지도

논문접수일 :2015년 8월 30일 논문수정일 :2015년 9월 7일 게재확정일 :2015년 9월 9일 투고유형 :
영문일반 교신저자 :신경식

* 이화여자대학교 경영대학

** 교신저자: 신경식

이화여자대학교 경영대학

120-750 서울특별시 서대문구 이화여대길 52 Tel: 02-3277-2799,

Fax: 02-3277-2776, E-mail: ksshin@ewha.ac.kr

하이브리드 인공 신경망 모델을 사용한 파산 유형 예측

저자 소개



조남옥

이화여자대학교에서 빅데이터 분석 기법을 경영분야에 적용하는 연구로 경영학 박사 학위를 취득하였고, 현재 이화여자대학교 경영대학 박사후 연구원으로 재직 중이다. 주요 연구분야는 지능형 의사결정지원시스템, 데이터 마이닝, 텍스트 마이닝, 빅데이터 분석 및 응용 등이다.



김현정

이화여자대학교 통계학사, 경영학 석/박사 학위를 취득하고, 한국산업은행 리스크관리본부와 딜로이트컨설팅에 재직했으며, 현재 이화여자대학교 경영대학 연구교수로 재직 중이다. 주요 연구분야는 지능형 의사결정지원시스템, 데이터 마이닝과 인공지능 응용, 빅데이터 분석 및 비즈니스 인텔리전스 등이다.



신경식

현재 이화여자대학교 경영대학 경영학부 교수로 재직 중이다. 연세대학교 경영학과를 졸업하고 미국 George Washington University에서 MBA, 한국과학기술원(KAIST)에서 인공지능, 지식기반 시스템 등 지능형 기법을 경영분야에 적용하는 연구로 경영공학 Ph.D.를 취득하였다. 주요 연구분야는 데이터 마이닝과 비즈니스 인텔리전스, 빅데이터 분석/비즈니스 애널리틱스, 인공지능 응용과 지식공학 등이다.