

부스팅 인공신경망을 활용한 부실예측모형의 성과개선

김 명 종¹, 강 대 기²

¹동서대학교 경영학부, 교신저자

²동서대학교 컴퓨터정보공학부

Boosting neural networks with an application to bankruptcy prediction

Myoung-Jong Kim^{1*}, Dae-Ki, Kang²

ABSTRACT

In a bankruptcy prediction model, the accuracy is one of crucial performance measures due to its significant economic impacts. Ensemble is one of widely used methods for improving the performance of classification and prediction models. Two popular ensemble methods, Bagging and Boosting, have been applied with great success to various machine learning problems using mostly decision trees as base classifiers. In this paper, we analyze the performance of boosted neural networks for improving the performance of traditional neural networks on bankruptcy prediction tasks. Experimental results on Korean firms indicated that the boosted neural networks showed the improved performance over traditional neural networks.

Keywords: Neural Networks; AdaBoost; Boosting Neural Networks; Bankruptcy Prediction.

1. 서 론

부스팅 알고리즘은 복수의 약분류자 앙상블 (weak classifiers ensemble)을 구성하고 앙상블의 예측 결과를 가중결합하여 최종 예측결과를 도출할 수 있는 단일 강분류자 (a single strong classifier)를 산출하는 기법으로 정의할 수 있다. 최근 부스팅 의사결정트리를 부실 예측에 적용한 연구들이 활발하게 진행되었다 (Alfrazo, Gámez, & García, 2007; Alfrazo, García, Gámez, & Elizondo, 2008). 유럽 기업을 표본대상으로 한 분석결과는 인공신경망과 비교하여 부스팅 의사결정트리는 오분류율을 약 30% 감소시키는 성과개선 효과가 있음을 보여주었다. 그러나, 부스팅 의사결정트리의 연구와 비교하여 부스팅 인공신경망 (Boosting neural networks)을 적용한 부실예측 연구는 현재까지 보고되지 않고 있다. 이는 부스팅 의사결정트리에 비교하여 나이브 베이즈 및 선형판별식 분석에 기초한 안정적인 알고리즘에서는 부스팅 알고리즘의 성과개선 효과가 상대적으로 크게 나타나지 않기 때문이다. 이러한 이유로 대부분의 부스팅 알고리즘은 의사결정트리의 앙상블에 초점을 맞추고 있으며 국내에서도 부스팅 의사결정트리

연구가 활발하게 진행되고 있다 (Seok & Ryu, 2002; Kim et al., 2005; 이영섭 외, 2005; Chung & Kim, 2005; 최성하 외, 2007).

본 연구에서는 인공신경망 부실예측 모형의 판별력 개선을 위한 부스팅 인공신경망의 효과를 분석하고자 한다. 부가적으로 다양한 인공신경망의 구조하에서 부스팅 기법의 성과개선 효과를 분석하기 위하여 서로 다른 구조를 가진 다양한 인공신경망에 부스팅 기법을 적용하여 성과 개선의 효과를 분석하고자 한다.

본 논문은 구성체계는 다음과 같다. 제 2장에서는 대표적 부스팅 알고리즘인 AdaBoost 알고리즘에 대하여 소개하고자 한다. 제 3장에서는 표본수집 및 재무비율 선정과 관련된 연구설계 과정과 본 연구에서 대체적 기법으로 활용되고 있는 인공신경망과 부스팅 인공신경망의 설계 프레임워크를 설명하고자 한다. 제 4장과 제 5장에서는 각각 주요 연구 결과와 향후 연구방향을 제시하고자 한다.

2. AdaBoost 알고리즘

AdaBoost알고리즘은 앙상블 학습 알고리즘 중 가장

일반적으로 사용되고 있는 부스팅 알고리즘으로 Freund & Schapire (1997)에 의하여 제안되었다. 부스팅은 임의 예측보다 다소 높은 수준의 정확성을 보유한 여러 개의 약분류자의 선형결합으로 정확성이 높은 강분류자를 생성하는 알고리즘이다. 부스팅 방법은 이전 분류자의 성과를 기초로 오분류된 관측치에 초점을 맞추어 분류자를 순차적으로 생성한다. AdaBoost 알고리즘의 설명을 위하여 n 개의 학습 표본과 K 개의 기저 분류자로 구성된 앙상블 $C = \{C_1, C_2, \dots, C_k\}$ 을 가정하면 k 번째 기저분류자의 오류율(e_k)은 다음과 같이 단순평균으로 계산된다.

$$e_k = \frac{1}{n} \sum_{i=1}^n L(C_k(x_i), y_i)$$

$$L(C_k(x_i), y_i) = \begin{cases} 1 & C_k(x_i) \neq y_i \\ 0 & C_k(x_i) = y_i \end{cases}$$

여기에서 x_i 는 i 번째 관측치의 예측변수 벡터이고 y_i 는 i 번째 관측치의 범주를 나타내며 $C_k(x_i)$ 는 예측변수 벡터 x_i 에 대한 k 번째 분류자의 분류결과이다. $k+1$ 번째 분류자에서 i 번째 관측치에 부여되는 가중치는 $w_{k+1}(i) = w_k(i) \exp(\alpha_k L(C_k(x_i), y_i))$ 로 조정되어 오분류된 관측치에 더 높은 가중치가 부여된다. 여기에서 α_k 는 분류자의 중요도 또는 정확도의 개념으로 해석되며 $\alpha_k = \ln((1 - e_k)/e_k)$ 로 계산된다. $k+1$ 번째 분류자의 학습표본을 구성할 때 가중치가 높은 오분류 관측치가 많이 포함되기 때문에 부스팅 알고리즘은 오분류 관측치에 초점을 맞춘 학습을 진행할 수 있게 된다. 앙상블

학습은 $e_k < 0.5$ 일 때 학습을 중단하며 i 번째 관측치의 최종결과를 다음과 같이 앙상블의 결과치의 가중평균으로 계산하여 산출한다.

$$C(x_i) = \text{sign} \left(\sum_{k=1}^K \alpha_k C_k(x_i) \right)$$

3. 연구 설계

부실 기업은 2002~2005년 중 은행연합회 신용정보 등록기업, 당좌부도 발생기업, 회사정리절차 개시기업, 기업구조조정절차 개시기업에 해당하는 729개 외부감사 제조기업으로 구성하였으며 부실 기업에 대한 재무자료는 부실 직전 년도를 중심으로 수집하되 직전년도 재무자료가 없는 경우 2년 전 재무자료를 수집하였다. 부실기업에 대응되는 정상 기업은 외부감사 제조기업 중 2005년 말 기준으로 부실 사유에 해당하지 않는 729개의 기업으로 구성하였다.

부실 예측에 사용되는 재무비율은 일차적으로 기존의 기업부실 예측연구에 사용된 비율 및 실무에서 부실예측의 지표로 유용하게 활용되는 비율을 중심으로 30개의 재무비율을 수집하였다. 수집된 재무비율을 수익성, 부채상환능력, 레버리지, 자본구조, 유동성, 활동성 및 규모의 7개 재무비율 군으로 재분류하였다. 최종 입력변수는 ROC 분석의 AUROC를 이용하여 각 분류군별로 AUROC가 높은 7개 재무비율을 선정하였다. 일차적으로 선정된 30개 비율 및 최종 선정된 7개 재무비율의 AUROC는 <표 1>에 제시되어있다.

표 1. 재무비율의 AUROC (* 최종 7개 재무비율)

분류군	재무비율	AUROC	분류군	재무비율	AUROC
수익성	총자산경상이익율*	52.1	레버리지	총부채/총자산 *	52.4
	총자산순이익율	45.7		자기자본비율	51.9
	금융비용/매출액	49.2		유동자산/총자산	51.3
	금융비용/총부채	48.5	자본구조	이익잉여금/총자산*	53.6
	순금융비용/매출액	50.3		이익잉여금/총부채	51.6
	매출액경상이익율	45.7	유동성	이익잉여금/유동자산	50.8
	매출액순이익율	49.9		현금비율*	46.5
	자기자본경상이익율	48.2	당좌비율	45.7	
자기자본순이익율	47.5	유동비율	43.2		
부채상환	EBITDA/이자비용 *	53.2	활동성	재고자산회전율 *	31.5
	EBIT/이자비용	49.2		유동부채회전율	28.5
	영업현금흐름/이자비용	48.5	규모	매출채권회전율	27.0
	영업현금흐름/총부채	47.8		총자산*	25.2
	영업현금흐름/이자비용	51.9	매출액	22.4	
	영업현금흐름/총부채	52.7	고정자산	24.6	
	부채상환계수	50.8			
	차입금/이자비용	52.4			

4. 연구 결과

본 연구에서 벤치마킹으로 사용될 인공신경망은 입력층에 7개의 입력노드와 출력층에 2개의 출력노드 및 1개의 은닉층을 가진 전형적인 다층 퍼셉트론 (Multi-Layer Perceptron: MLP)이다. 인공신경망의 구조의 다양한 변화에 따른 부스팅 인공신경망의 성과개선의 안정성을 분석하기 위하여 은닉층의 은닉노드의 수를 5개~14개로 다양하게 구성하여 10개의 서로 다른 구조를 가진 인공신경망을 구성하였다.

본 연구에서는 AdaBoos 알고리즘을 활용하여 부스팅 인공신경망을 구성하였다. Opitz와 Maclin(1999)에 의하면 부스팅은 분류자 생성의 반복회수가 25번이 될 때까지 오분류율이 개선이 되었지만 그 이후에는 오분류율의 감소에 별 영향을 받지 않는다고 제안하였다. 이러한 연구결과를 기초로 본 연구에서는 부스팅 인공신경망의 분류자 생성 회수를 25회로 제한하였다.

본 연구에서는 인공신경망과 부스팅 인공신경망의 성과차이가 우연한 결과가 아님을 확인하기 위하여 10-fold 검증을 5회 반복 수행하였다. 이를 위하여 전체 1458개의 기업을 표본수가 동일한 10개 fold로 구성하고 9개의 집합은 분석용 데이터로 활용하고 나머지 1개 fold를 검증용 데이터로 활용하게 된다. 이러한 방법으로 50회의 교차타당성 검증을 수행하였다.

<표 2>는 50회 교차타당성 검증결과로서 학습용 데이터와 검증용 데이터에 대한 인공신경망과 부스팅 인공신경망의 평균 정확도를 보여주고 있다. 성과 비교 측면에서 부스팅 인공신경망은 전통적 인공신경망에 비하여 10개의 모든 신경망구조에서 약 3% 이상의 정확성 개선을 일관성있게 보여주고 있으며 이러한 결과는 부스팅 인공신경망의 성과개선의 효과가 안정적인 것을 의미한다. 50개의 실험 결과를 대상으로 t-test 분석

을 실행한 결과 부스팅 인공신경망은 1% 유의수준에서 인공신경망보다 높은 정확성을 보여주었다.

5. 결론 및 향후 연구 방향

본 연구에서는 부실예측 인공신경망의 성과개선을 위하여 부스팅 기법과 인공신경망이 결합된 부스팅 인공신경망을 제안하였다. 실험결과들은 부스팅 인공신경망은 10개의 다양한 인공신경망의 성과를 안정적으로 개선하였고 과도적합에 대한 견고성을 강화하였음을 보여주고 있다. 또한 전체 오분류율을 8.60% 감소시키며 제 1종 오류를 76.72% 감소시킬 수 있음을 보여주었다. 이는 부스팅 인공신경망이 부실예측문제에 있어 안정적으로 적용될 수 있음을 의미한다.

그러나 본 연구의 결과를 다양한 경영분야에 적용하기 위하여 향후 다음과 같은 연구를 진행하고자 한다. 첫째, 본 연구에서 활용된 AdaBoost.M1 이외에도 다양한 부스팅 알고리즘이 제안되었다. 예를 들어 Confidence rated boosting (Schapire & Singer, 1999), Margin Boost (Mason et al., 2000), Logit Boost (Friedman, 2001) 등 다양한 부스팅 알고리즘의 적용을 통하여 도메인별 최적 알고리즘에 대한 탐색과정이 필요하다.

둘째, 본 연구의 결과를 효과적으로 활용하기 위해서는 데이터 특성에 따른 부스팅 기법의 효과와 같은 심층 분석이 추가되어야 한다. 예를 들어 데이터에 내재된 노이즈 (noise)가 학습결과에 미치는 영향에 대한 분석을 통하여 다양한 분류기법과 부스팅 기법의 결합에 대한 연구가 더욱 활발해질 수 있을 것으로 기대된다.

표 2. 예측정확성의 비교 (단위: %)

구조	인공신경망			부스팅 인공신경망		
	학습표본	검증표본	성과차이	학습표본	검증표본	성과차이
7 5 2	74.71	71.06	3.66	75.58	75.04	0.54
7 6 2	74.81	70.78	4.03	75.29	75.03	0.26
7 7 2	74.73	71.12	3.60	75.51	75.24	0.28
7 8 2	74.84	71.12	3.71	75.95	74.90	1.05
7 9 2	74.81	70.92	3.89	75.92	75.24	0.67
7 10 2	74.82	70.92	3.90	75.54	74.62	0.91
7 11 2	74.84	70.85	3.99	75.76	75.58	0.17
7 12 2	74.80	71.19	3.60	75.66	74.96	0.70
7 13 2	74.84	71.12	3.72	75.87	75.18	0.70
7 14 2	74.84	71.12	3.30	75.96	75.18	0.78
평균	74.80	71.02	3.78	75.70	75.10	0.60

참고문헌

- [1] Alfaro, E., Gámez, M., & García, N. (2007). Multiclass corporate failure prediction by AdaBoost.M1. *Advanced Economic Research*, 13, 301-312.
- [2] Alfaro, E., García, N., Gámez, M., & Elizondo, D. (2008). Bankruptcy forecasting: an empirical comparison of AdaBoost and neural networks. *Decision Support Systems*, 45, 110-122.
- [3] Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, 36, 105-139.
- [4] Drucker, H., & Cortes, C. (1996). Boosting decision trees, *Advanced Neural Information Processing Systems*, 8.
- [5] Freund, Y. (1995). Boosting a weak learning algorithm by majority. *Information and Computation*, 121(2), 256-285.
- [6] Freund, Y., & Schapire, R. E. (1996). Experiments with a new boosting algorithm. *Machine Learning: Proceedings of Thirteenth International Conference* (pp. 148-156).
- [7] Freund, Y., & Schapire, R. E. (1997). A decision theoretic generalization of online learning and an application to boosting. *Journal of Computer and System Science*, 55(1), 119-139.
- [8] Friedman, J. (2001). Greedy function approximation: a gradient boosting machine, *The Annals of Statistic*, 29(5), 1189-1232.
- [9] Maclin, R., & Opitz, D. (1997). An empirical evaluation of bagging and boosting. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, (pp. 546-551).
- [10] Mason, L., Baxter, J., Bartlett, P., & Frean, M. (2000). Functional gradient techniques for combining hypotheses. In A. J. smola, P.L. Bartlett, B. ScholkDirect optimization of margins improves generalization in combined classifiers. In M. S. Kearns, S. Solla, & D. Cohn (Eds.), *Advances in Neural Information Processing Systems*, 11, Cambridge, MA: MIT Press.
- [11] Pantalone, C., & Platt, M. B. (1987). Predicting commercial bank failure since deregulation. *New England Economic Review*, 37-47.
- [12] Optiz, D., & Maclin, R. (1999). Popular ensemble methods: an empirical study. *Journal of Artificial Intelligence*, 11, 169-198.
- [13] Quinlan, J. R. (1996). Bagging, boosting and C4.5. *Machine Learning: Proceedings of the Fourteenth International Conference* (pp. 725-730).
- [14] Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197-227.
- [15] Schapire, R. E., & Singer, Y. (1999). Improved boosting algorithm using confidence-rated predictions, *Machine Learning*, 37(3), 297-336.