

SVM-NN Hybrid모형을 활용한 호텔도산예측*

Cost-conscious SVM-NN Hybrid Model for the Hotel Bankruptcy Prediction

金秀英**

Kim, Soo-Young

ABSTRACT

This study proposes an integration strategy regarding the efficient prediction of hotel bankruptcy by combining data mining techniques. In particular, by combining support vector machine(SVM) and neural network(NN), SVM based NN hybrid model for hotel bankruptcy prediction is newly introduced in this study. In the experiments on Korea deluxe hotel data, SVM-NN hybrid model achieves a performance accuracy of 96.34%, which is better than that of stand-alone classifiers. The hybrid model performs better in the grey area where some bankrupt hotels appear to be less financially distressed. The results suggest that debt-burdened hotels with low profit margin and ordinary income margin as well as lower growth in asset are more likely to be candidates of bankruptcy. Accurate bankruptcy prediction usually brings into many benefits such as risk reduction in investment return, better monitoring, and an increase in profit. Limitations of the study and avenue for future research directions are also discussed at the end.

핵심용어(Key words) : 호텔도산예측(Hotel bankruptcy prediction),
SVM-NN hybrid 모형(Model),
예측정확성(Performance accuracy)

www.kci.go.kr

* 이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2009-327-B00879).

** 세종사이버대학교 호텔관광경영학부 부교수. e-mail: sookim.korea@gmail.com

I. 서론

주5일 근무제 실시, 소득수준 향상 등에 따른 레저 활동 인구증가로 휴양콘도 운영업 및 기타 관광 숙박시설운영업의 사업체수가 2001년 대비, 크게 증가하였음에도 불구하고, 2005년의 호텔 사업체수와 영업이익율은 국내관광산업의 부진으로 2001년 대비, 각각 16.6%와 2.9%가 감소하였다(통계청, 2007). 2011년 통계청의 “2004-2009 사업체 생성·소멸 현황 분석”에 의하면 숙박 및 음식점의 신규사업체수가 124,299로 전체산업의 20.88%를 차지하는 동시에 휴폐업사업체 수 역시 127,443로 22.07%를 차지하여 도매 및 소매업을 제외하고는 최고의 휴폐업사업체 수치고 비율을 차지하는 것으로 나타났다. 한국은행 경제통계국 기업통계팀의 2007년부터 2010년까지의 평균자료에 의하면, 제조업과 서비스업에 비교할 때, 숙박업의 총자산 순이익율과 자기자본순이익율은 각각 -0.70배, -0.55배와 -0.75배, -0.66배 만큼 상대적으로 낮은 반면 금융비용대매출액은 오히려 2.26배, 1.84배 만큼 높게 나타나, 숙박업의 손익관계비율이 타 산업에 비교하여 열세를 보이고 있다(<http://ecos.bok.or.kr/>, 2011.9.1). 이렇듯 최근의 사업 환경 악화와 맞물려 호텔산업의 특성을 반영한 정확한 도산예측이 필요함에도 불구하고(Brigham & Gapenski, 1994), 지금까지 진행된 국내호텔 도산예측 및 분석에 대한 연구는 극소수이다(김수영, 2006; 김수정 & 김우곤, 1998; 김시중, 2005).

기업도산예측은 Altman(1968)이후 회계와 재무영역에서 주요 주제로 사용되어왔는데, 이는 기업도산이 금융권의 대출의사결정과 투자자의 투자여부결정, 채권자나 소유자의 수익성 확보 등에 주요한 영향을 미치기 때문이다. 분류 방법론적 접근으로는 1980년대 초반까지 다변량판별분석(multivariate discriminant analysis: MDA)과 logistic 회귀분석방법 등의 통계적인 방법을 주로 사용하였으나(Dimitras, Zanakis & Zopounidis, 1996), 이 두 가지 방법의 보편성이 예측정확성에 근거한다는 증거는 확실하지 않다. 1980년대 중반 이후에는 주식시장예측 및 도산예측 등에 인공지능(artificial intelligence: AI) 방법을 사용해 왔으며(Braun & Chandler, 1987; Lee & Oh, 1990), 특히 경영예측 및 채권등급결정, 주식가격예측, 이자율예측 등에 Neural Networks (NN)를 많이 활용하고 있다(Jhee & Lee, 1993; Kim & Noh, 1997; Surkan & Singleton, 1990). 연속적 시간자료를 분석하는 연구의 경우에는 특히 NN에 의한 분류가 기타의 통계학적 방법보다 더 정확하게 나타나는데, 이는 NN이 데이터 세트의 비선

형관계를 규명하고 실험하는데 더 적합할 뿐만 아니라 다변량 표본분포 혹은 사전 확률조건과 관련된 전제를 필요로 하지 않기 때문이다(Azia & Dar, 2006; Ohlson, 1980; Hammer, 1983; Wong, Bondnovich, & Yakup, 1997).

또한 최근 들어 Vapnik(1995)가 개발한 Support Vector Machine(SVM)을 도산예측에 많이 활용하기 시작하였는데, 이는 SVM모형이 이론적으로 잘 무장된 전통적인 통계방법과 자료특성을 정확히 반영하면서 분포에 대한 전제조건으로부터 자유로운 기계학습 시스템(machine learning system)을 잘 통합하기 때문이다(Ravi Kurma & Ravi, 2007). SVM은 decision tree 및 NN, Bayesian 방법 등의 기타 알고리즘 보다 분류정확도가 우수한 것으로 인정받고 있다(Joachims, 1998; Li, Sun & Sun, 2009; Min & Lee, 2005; Min, Lee & Han, 2006; Mukkamala, Janowski & Sung, 2002; Pal & Mather, 2004; Ramaswamy, Tamayo, Rfkin, Mukkherjee, Yeang, Angelo, Ladd, Reich, Latulippe, Mesiroy, Poggio, Gerald, Loda, Lander, & Golub, 2001; Shin, Lee & Kim, 2005; Wu, Tzeng, Goo, & Fang, 2007).

현재까지 하스피탈리티산업에 대한 도산예측연구는 다양한 분류방법들을 독립적으로 활용하거나 이들 간의 비교에 초점을 맞추어왔는데(Davalos, Gritta, & Chow, 1999; Gao, 1999; Gu, 2002; Hanson, 2003; Kim & Gu, 2006; Lipovats, Mandaaraka & Mourelatos, 2000; Moncars & Kron, 1993; Noronha & Singal, 2004; Parsa, Self, Njite & King, 2005; Patterson, 1971; 김수영, 2006), 이러한 단독의 모형들은 기타의 분류 방법들이 제공하는 차별적인 정보들을 복합할 수 없다는 단점을 갖는다(Kittler, 1998). Zhou & Lopresti (1997)는 반복적 표본을 사용하여 훈련(training)시킨 통합적인 기계 학습 모형이 분류정확성이 향상됨을 확인하였으며, 이는 학습모형의 hybrid가 단일의 최고 모형보다 예측정확성을 높일 수 있음을 의미한다. 최근 몇 년간의 부도 예측은 단독 분류방법간의 통합을 모색하여왔으며 많은 연구결과를 통해 이러한 방법이 예측정확성을 개선한다는 점을 확인하였다(Asaad, Bone & Cardot, 2008; Cho, Kim & Bae, 2009; Esteban, Garcia, Gamez & Elizondo, 2007; Hua *et al.*, 2007; Kittler, 1998; Lee *et al.*, 1996; Maqsood, Khan, & Abraham, 2004; Optiz & Maclin, 1999; West, Dellana, & Qian, 2005; Yim & Michelle, 2002).

본 연구는 김수영(2006)과 Kim(2011)의 후속연구로 hybrid모형개발이라는 접근방법을 통해 도산예측과 관련된 변수를 확산하고 새로운 모형을 제시함으로써

호텔도산예측에 대한 이론적 이해를 증진하는 동시에 최적의 호텔도산예측모형을 개발하는데 그 목적이 있다. 주요 재무비율을 사용하여 이미 그 예측 우수성이 검증된 NN(김수영, 2006)과 최근에 도산연구에서 각광을 받고 있는 SVM(Kim, 2011, Min *et al.*, 2006)을 결합한 모형을 구축함으로써 오분류로 인한 경험적 위험을 최소화할 수 있는 최적의 호텔도산예측 모형을 제공하는데서 그 기여도를 찾을 수 있다. 연구결과, 두 가지 이상의 모형을 혼합한 hybrid 모형이 각 모형의 장점을 복합적으로 고려하기 때문에 McKee & Lensberg(2002)의 연구결과와 같이 예측정확성을 개선시킨다는 것을 확인하였다.

II. 이론적 배경

하스피탈리티산업에 대한 도산예측연구는 김수영(2006)에서 비교적 자세히 언급하였으므로 그 이후에 이루어진 연구를 보면 아래와 같다. 김수영(2006)은 8개 국내 특급호텔의 도산 전 3년 자료를 근거로 MDA, logistic, NN의 호텔도산 예측모형을 개발하였다. 예측결과 세 가지 모형이 각각 70%, 80%, 73%의 예측정확성을 보이며, 유동비율과 당좌비율, 매출액순이익율, 자기자본증가율, 고정장기 적합율이 호텔도산예측에 유의적인 변수인 것을 확인하였다. Kim(2011)은 33개 국내 관광호텔의 재무제표를 활용하여 MDA, logistic, NN, SVM을 활용한 호텔도산예측모형을 개발하였으며, 각각의 모형이 72.6%, 80%, 91.6%, 95.95%의 예측정확성을 보였다. 본 연구는 위의 두 선행연구를 기반으로 SVM-NN hybrid모형이라는 새로운 방법론을 적용하여 도산예측모형의 예측정확성을 높이고 이를 통해 gray area에 있는 호텔의 도산예측에 대한 추가변수를 찾고자 노력하였다. Youn & Gu(2011)는 Chapter 11에 의해 도산처리된 미국 레스토랑의 재무자료를 근거로 Logistic과 NN에 의한 도산예측모형을 구축하였으며, 각각의 모형은 모두 최고수치를 보이는 도산 2년 전 예측모형을 기준으로, 88.1%의 예측정확성을 보였다. 단, 10개의 재무비율에 근거하여 개발된 Youn과 Gu의 예측모형에서는 NN에서는 이자보상배율, logistic의 경우는 총부채에 대한 EBITDA만이 유의적으로 나타나 현실 적용에 한계를 갖는 것으로 판단된다. 허양희·설훈구·홍재범(2010)은 외부감사 존속 여부에 의해 부실을 정의하여 41개 호텔을 대상으로 15개 재무비율과 비재무적 요인을 변수로 로지스틱모형을 구축하였다. 이들의 예측모형은 호텔기업의 외부차입이 부실에 가장 유의적인 영향을 미치는

변수임을 증명하였으나, 3년간 외부감사를 받지 않는 기업을 부실기업으로 정의한 것에 의한 한계를 가지고 있다.

Hybrid 방법의 기본적인 개념은 일련의 분류방법을 통합함으로써 단독 분류방법의 장점을 통합한다는 것이다. 대표적인 통합방법인 Voting이나 Stacking 등을 활용하여 두 가지 이상의 분류방법을 통합함으로써 개별 분류방법의 예측성을 향상시킨 연구사례를 <표1>에 제시하였다.

두 가지 Voting 방법인 Bagging과 Bboosting은 상이한 양상의 임의적 하부 세트를 훈련함으로써 Base 분류방법에 다양성을 부여한다(Breiman, 1996b; Freund & Schapire, 1996; Schapire, 1990). Bagging은 원래의 훈련 세트로부터 다수의 Bootstrap 훈련 세트를 만들고 이를 복합적인 모형에 포함시킴으로써 새로운 분류양상을 창출한다(Breiman, 1996a; Dietterich, 2000; Efron & Gong, 1983). Optiz & Maclin (1999)은 Bagging이 단일 분류방법보다 거의 늘 예측력이 높다는 것을 증명하였다. 반면에, Bagging의 기능은 hybrid 분류방법에 단지 훈련 데이터를 제시하는 것에 그치기 때문에 Base 분류방법의 선택에 따라 예측력이 상이할 수 있다는 지적도 있다(Yang & Browne, 2004).

Boosting은 모형을 응집시키는 방법(Hastie, Tibshirani & Friedman, 2001)으로 훈련 세트에 대해 일련의 상이한 가중치를 부여한다(Freund & Schapire, 1996). Boosting 알고리즘은 우선 훈련세트에 초기 분포를 결정하고, 두 번째로 적응적 가중치(adaptive weight)에 의해 과정이 중단될 때까지 훈련을 지속한다. 그리고 마지막으로 훈련과정이 끝나면 훈련 세트에 근거해 가장 정확한 최종의 분류방법에 지금까지 획득한 단독 분류방법들을 결합시킨다(Asaad *et al.*, 2008). 따라서 최종의 분류방법은 일반적으로 높은 예측력을 갖게 된다(Asaad *et al.*, 2008; Esteban *et al.*, 2007; Maqsood *et al.*, 2004; West *et al.*, 2005). 반면에 Freund & Schapire(1996)는 boosting 이 훈련 세트의 overfittng으로 인한 예측력의 손상에 우려를 표하였다.

Stacking 혹은 훈련에 의한 결합방법은 Base Learner위에 두 번째 학습 기제를 겹쳐서 사용하는 방법이다(Duin, & Tax, 2000; Wolpert, 1992). 이 방법은 Base Learner의 산출물(output)을 중간공간의 매체로 활용하여 두 번째 단계의 기제인 Meta 분류단계에 재투입함으로써 Base Learner의 훈련된 산출물과의 결합을 꾀한다(Valentini & Masulli, 2002). Stacking은 단독 분류방법들을 복합하는데 유용할 뿐 아니라 단독 분류방법의 예측력을 개선하는데도 사용되고 있다(Ulas, 2009). 기본적으로 Meta 분류방법은 Base 분류방법의 산출물이 겪은

오류를 배우기 위한 훈련과정을 겪는다. Stacking은 모형에 포함된 단독분류방법들이 수정된 학습패턴을 통해 학습할 수 있기 때문에 분류 정확성이 개선되는 것으로 증명되었다(Charalambous, Charitous, & Kaourou, 2000; Cho *et al.*, 2009; Hua *et al.*, 2007; Kittler, 1998; Lee *et al.*, 1996; Yim & Mitchell, 2002). Stacking에 의한 분류방법은 Base 분류방법과 meta 분류방법을 어떻게 적절하게 구성할 것인가 하는 문제를 가지고 있는데, 만약 분류방법과 알고리즘의 수가 많지 않다면 지속적인 시도를 통해 최적의 결합을 구성할 수 있다(Ledezma, Aler, Sanchis, & Borrajo, 2010). Stacking 방법의 효율성은 base learner의 오류 독립성에 근거한다(Valentini & Masulli, 2002).

〈표 1〉 도산예측에 사용된 Hybrid 연구 사례

기존연구	우세한 hybrid 모형	비교모형
Chen, Ma, & Ma(2009)	SVM+MARS	SVM+CART, SVM, MARS, CART
Hua <i>et al.</i> (2007)	SVM+logistic regression	SVM
Min <i>et al.</i> (2006)	SVM-GA	NN
Assaad <i>et al.</i> (2008)	NN-Boosting	Regressions
Lee <i>et al.</i> (1996)	NN-SOFM > NN-ID3 > NN-MDA	MDA, ID3
Pendharkar(2001)	GA-NN	NN
Schmitz, Aldrich, & Gouws(1999)	NN-DT	CART
Kittler(1998)	GN-NN	NN
Cho <i>et al.</i> (2009)	Integration strategy with subject weight	MDA, logit, NN, DT
Maqsood <i>et al.</i> (2004)	Boosting, bagging, arching NN	MLPN, ERNN, RBFN, HFM
Esteban <i>et al.</i> (2007)	AdaBoost	NN
Olmeda & Fernandex(1997)	Simple voting	NN, C4.5, logit, MARS, MDA
Hung & Chen(2009)	Selective ensemble	Voting ensemble, Weighting ensemble, DT, BPN, SVM
Ledezma <i>et al.</i> (2010)	GA-Stacking	>Boosting > Bagging

주: 기존의 연구자료를 근거로 저자 작성.

SVM은 Hybrid모형의 Base 분류방법으로 많이 사용되어왔다(Verwoerd & Hunt, 2002; Wang, Zhao, Tan & Zhang, 2008; Wu, 2011). 아래와 같은 장점에 근거하여 본 연구 역시 SVM을 base분류방법으로 사용하고자 한다. 첫째,

SVM은 경험적으로 위험을 최소화하기 보다는 구조적으로 위험을 최소화하는 원칙에 근거하기 때문에 Overfitting의 문제를 극복할 수 있다(Min & Lee, 2005). 둘째, SVM은 분류방법의 복잡성을 모형화할 수 있는 Parameter c 를 활용하여 이상치를 무시함으로써 분류방법에 융통성을 부여한다(Doran, Raicu, Furst, Settimi, Schipma & Chandler, 2007). 셋째, SVM은 어떠한 자료에 서도 차별화된 자료를 찾아내는데 보다 효율적이다(Kittler, 2000). 넷째, SVM은 표본수가 지나치게 많지 않은 경우 기타의 예측기술보다 예측력이 뛰어나다(Shin *et al.*, 2005). 그리고 마지막으로 SVM은 기타의 분류방법과 복합적으로 사용할 경우 예측정확성을 높이는 것으로 증명되었다(Mei, Shen & Ye, 2008).

연구자는 본 논문을 통해 stacking방법을 활용한 SVM에 근거한 neural network(NN) hybrid모형을 구축하여 국내 호텔도산을 예측하였다. SVM-NN hybrid모형의 가장 큰 장점은 독립성 및 동질적 분포와 관련된 제한적 가정으로부터도 자유로울 수 있기 때문에 보다 광범위한 업계의 활용이 가능하다는 점이다.

III. 연구방법

1. 표본설계

본 연구에서의 도산호텔의 정의는 아래의 다섯 가지 요건 중 하나 이상을 만족할 경우로 정의하였다: (1) 경제적 실패를 의미하는 채무불이행 호텔(Beaver, 1966; Lev, 1969); (2) 기술적 지급불능의 상태로 재무적 의무에 도달했을 때 이를 지불할 수 없어 부채 재조정을 시도한 호텔(Tavlin, Moncrz & Dumont, 1989); (3) 청산 등의 이유로 금융감독원 등록을 취소한 호텔(Dun & Bradstreet, 1994); (4) 폐업으로 사업을 더 이상 진행하지 않는 호텔(Booth, 1983; Deakin, 1972); (5) 도산과정 중 제3의 기업에게 인수 합병된 호텔(Tavlin *et al.*, 1989).

많은 연구들이 자료수집과정에 있어 도산기업과 비도산기업의 차이를 확인하기 위해 자산이나 자본규모, 매출액(Fletcher & Goss, 1993; Leshno & Sharda, 1990; Nanda & Pendharkar, 2001; Tsukuda & Baba, 1994), 산업분류나 경제 부문(Raghunpathi, Schkade & Raju, 1991), 지리적 위치(Rumelhart, Hinton & Williams, 1986), 또는 사업장의 수나 연한, Charter

Status(Tam & Kiang, 1992) 등을 기준으로 한 쌍대표본을 사용해왔다. 본 연구는 한국신용정보(주)와 금융감독원에서 제공하는 국내 관광호텔들의 재무제표를 취합한 후 재무요소들을 사용하여 Hybrid 호텔도산예측모형을 개발하였다. 도산에 이르는 호텔들의 재무적 특성을 확인하기 위해 1995년부터 2003년 기간 중에 도산한 호텔을 기준으로 총자산의 규모가 같은 호텔들을 쌍대표본으로 하여, 도산이전 12개월을 t-1년으로 정의한 후, 자료 확보여부에 따라 도산호텔과 비도산호텔의 재무자료를 t-1부터 t-3까지로 대응하여 비교 분석하였다. 도산 3년 전의 재무자료를 포함한 18개 호텔과 2년 전의 자료를 포함한 12개 호텔, 1년 전의 자료를 포함한 3개의 호텔을 포함하였으며, 경제 불황이었던 1997년부터 1998년에 도산한 호텔이 전체 표본의 75.8%를 차지하고 있다. 최종표본은 33개의 도산호텔과 비도산호텔을 매칭하였으며, t-1에서 t-3에 이르는 기간 동안의 총 164개 표본으로 구성하였다. 또한 본 연구는 Clementine과 SPSS를 사용하여 호텔도산예측 모형을 구축하였다.

2. SVM-NN hybrid 분석

SVM은 비교적 새로운 기계학습 기술로 NN의 주요문제인 국소최저치(local minima)와 Overfitting의 문제를 해결하기 위해 개발되었다(Cristianini & Shawe-Taylor, 2000; Min & Lee, 2005; Vapnik, 1995; Vapnik, 1998; Kim, 2003). SVM의 기본적인 목적은 특성들을 보다 광범위한 특성공간으로 지도화하여 데이터들을 새로운 공간에서 연속적으로 분리한 후, 차이(margin)를 최대화하고 경험적 위험요소를 최소화할 수 있는 연속적인 분리 hyperlane을 찾아내는 것이다(Cristianini & Shawe-Taylor, 2000). SVM의 장점은 분리하기 힘들었던 데이터 세트를 kernel 기능을 활용하여 보다 높은 차원의 공간으로 지도화할 수 있다는 것이며(Kittler, 2000), 일반적으로 사용하는 kernel기능으로는 Polynomial Kernel과 Radial Basis Function이 있다(Vapnick, 1998). 또한 이상치를 무시할 수 있는 Parameter를 활용하여 두 계층을 분리할 수 있는 Hyperlane을 찾는데, 실험자 입장에서는 새로운 데이터세트를 축적해 가는 과정에서 이상치를 제거하기 위한 별도의 과정을 생략할 수 있다는 측면이 장점으로 다가올 수 있다(Doran *et al.*, 2007). 또한 SVM은 현실에 적용할 경우, 표본의 수가 지나치게 크지 않을 경우 한층 개선된 예측력을 갖는 것으로 증명되었다(Shin *et al.*, 2005). NN은 인공지능에서 독보적인 연구방법론으로 자리를 잡

았으며 도산예측을 포함한 분류연구에서 많은 연구자들이 NN을 활발하게 사용하고 있다. 많은 연구결과는 NN이 기존의 전통적인 통계학적 분류방법과 귀납적인 방법을 능가하는 예측결과를 보이며, 특히 독립변수와 종속변수간의 관계가 알려지지 않았거나 복잡한 경우에는 유용한 것으로 확인하였다(Boritz, Kennedy, & Albuquerque, 1995; Davalos *et al.*, 1999; Etheridge & Sriram, 1997; Huang, Chen, Hsu, Chen & Wu, 2004; Jo & Han, 1995; Koh, 2004; Lin & McClean, 2001; Odom & Shara, 1993; Raghupathi *et al.*, 1991; Swicegood & Clark, 2001; Tam & Kiang, 1992). NN이 가지고 있는 내재적인 단점은 어떠한 경로에 의해 결과물이 산출되었는지를 알수 없는 내부적 구조에 있으며, 이로 인해 명확한 설명력에는 한계가 있다. 그러나 연구자가 모형으로부터 정확한 분류결과 만을 원한다면 심각한 문제는 아니며, Clementine은 NN모형 내 입력노드의 상대적 중요성을 제공함으로써 이러한 단점을 극복하고자 하는 노력을 기울이고 있다. NN의 장점을 살펴보면, 첫째 결측 데이터나 불완전한 데이터를 잘 다룰 수 있다(Shachumurove, 2002). 둘째, NN은 데이터에 대해 통계분류방법에서와 같은 확률분포에 대한 기본가정을 요구하지 않는다(Sung, Chang & Lee, 1999). 셋째, 비교적 짧은 시간에 방대한 양의 정보를 처리할 수 있는 능력을 가지고 있다(Yoon, Guimaraes & Swales, 1994). 넷째, NN은 연속형 확률변수와 이산형 확률변수를 모두 다룰 수 있으며, 여러 가지 패키지를 통해 쉽게 다룰 수 있다(Goh, 1995; Lahiri, 2006). 그리고 마지막으로, 급속한 환경의 변화에 비교적 쉽게 모형을 적응시킬 수 있다(Sung *et al.*, 1999).

본 연구는 하스피탈리티 연구영역에서는 처음으로 Mei *et al.*(2008)과 Rong & Yuan (2010), Zhang, Berg, Maire & Malik (2006)이 의학 및 visualizing연구에 시도했던 방법인 SVM-NN hybrid모형을 사용하여 국내 호텔에 대한 도산을 예측하였다. 첫 번째의 단계 0에서는 SVM모형을 활용하여 잠재적으로 중요한 설명변수 세트를 규명하였으며, 두 번째의 단계 1에서는 NN을 활용하여 예측변수의 stacking을 모색하였다(Cho, Ahn, & Lee, 2001; Jiang, Zhou & Chen, 2002). 이러한 SVM-NN hybrid 접근방법의 사용은 두 가지 이유에 근거한다. 첫째는 SVM이나 NN은 확률분포에 대한 기본가정으로부터 자유롭기 때문에 그것을 무시함으로써 발생할 수 있는 내재적 문제점을 극복하고 예측력을 개선할 수 있다는 것인데, 이는 기존 연구자들이 실험적인 도산예측이 이러한 가정을 무시함으로써 내재적인 문제를 포함하게 된다는 것을 지적해왔기 때문이다(Dragos *et al.*, 2008; Grice & Dugan, 2001; Lennox,

1999; Zmijewski, 1984). 둘째는 hybrid모형을 사용함으로써 SVM과 NN이 제공하는 차별적인 정보들을 복합할 수 있다는 것이다.

IV. 분석결과

1. 기초통계량 분석

두 그룹 간의 재무비율의 차이를 확인하기위해 기술통계분석을 실시한 결과를 <표 2>에 제시하였으며 대부분의 변수들이 최소 0.10의 유의수준에서 유의적인 상이함을 보이고 있다.

<표 2> 도산호텔과 비도산호텔 간의 재무비율 비교

	재무비율	도산호텔	비도산호텔	t-statistic	p-value
유동성비율	유동비율	32.067	61.353	4.601	0.033 **
	당좌비율	29.004	51.308	2.941	0.088 *
	매출채권회전율	76.095	34.627	11.848	0.001 ***
안정성비율	부채비율	3029.627	813.218	22.778	0.000 ***
	고정장기적합율	608.333	283.278	4.290	0.040 **
수익성비율	매출액순이익률	-124.276	1.351	4.683	0.032 **
	총자본경상이익률	-306.162	26.081	4.382	0.038 **
	자기자본순이익률	-365.136	35.745	9.352	0.003 **
	자기자본경상이익률	-323.070	49.755	12.468	0.001 ***
활동성비율	총자본회전율	2.117	1.973	1.630	0.204
	재고자산회전율	58.888	61.723	0.037	0.847
	고정자산회전율	1.676	0.820	16.964	0.000 ***
성장성비율	매출액증가율	2.748	125.412	3.335	0.070 *
	총자산증가율	7.341	15.650	0.450	0.504
	경상이익증가율	154.739	19.188	5.814	0.017 **
	당기순이익증가율	14.868	-104.397	0.054	0.871
	자기자본증가율	6.926	66.888	8.074	0.005 **

주: ***p<0.01. **p<0.05, * p<0.10.

낮은 유동성과 안정성이 도산호텔의 수익성에 영향을 미치는 것으로 확인되고 있으며, 재고자산회전율을 제외한 활동성비율은 오히려 도산호텔이 높은 수치를 보이고 있다. 반면에 성장성비율은 매출액증가율과 자기자본증가율은 비도산 호텔이 높은 반면, 경상이익증가율은 도산호텔이 유의적으로 높은 것으로 나타나 이후 구체적인 Hybrid분석모형에 의한 도산원인분석이 필요하다.

2. SVM-NN Hybrid모형 분석

우선 단계 0에서는 잠재적으로 유의적인 설명변수들을 규명하기 위한 Baseline으로 예측력이 가장 정확하다고 알려진(Min *et al.*, 2006; Shin *et al.*, 2005; Wu *et al.*, 2007) SVM모형을 활용하였다. Linear와 polynomial, radial basis, sigmoid의 네 가지 kernel 기능을 0.0001의 Stopping criteria에 적용하여 개별적으로 테스트하였으며 가장 우수한 예측력을 보인 Third-degree Polynomial Kernel 기능을 선택하였다.

다음의 단계 1에서는 단계 0의 SVM Vector위에 예측변수를 Stack하기위해 NN을 사용하였다. 단계 1의 알고리즘으로 Quick과 Dynamic, Multiple, Prune, RBFN, Exhaustive Prune의 다양한 방법을 사용하여, 네트워크 성과에 더 이상의 증가가 없을 때까지 NN 패러다임을 테스트하였다. Overfitting을 방지하기위해 'Prevent overtraining parameter'를 사용하여 최상의 모형을 찾기 위한 노력을 기울였다. 최종적으로 Quick 알고리즘을 선택하였으며, 다양한 NN을 구축하고 테스트한 결과 Single-layer perception neural network이 가장 우수한 예측력을 갖는 것으로 확인하였다. Hidden Layer내의 Hidden Neurons의 수 역시 NN의 테스트 세트 결과에 근거하여 경험적으로 선택하였으며, 최종적으로 세 개의 Hidden Unit을 가진 단일 Layer Network의 Back Propagation 모형을 호텔도산예측 모형으로 구축하였다. Cybenko(1989)와 Hornik, Stinchcombe & White(1989)는 단일 Hidden Layer로 구성된 Network 구조가 적정한 예측력수준 내에서 복잡한 시스템을 모형화하는데 아무런 문제가 없다는 것을 확인하였다.

SVM-NN Hybrid 모형구축의 결과를 <표 3>에 제시하였으며, SVM-NN Hybrid모형은 도산호텔에 대해서는 100% 정확하게 예측하여 Type I 오류율이 0%이고, Type II 오류율은 7.32%이다. 전체적인 모형의 예측정확성은 96.34%로 우수한 예측성과를 보여주고 있다. 이러한 예측력은 Jo, Han & Lee(1997)의 평균 89.36%와 Sung *et al.*(1999)의 81-83.3%, Tam & Kiang(1992)의 65.2-94%, 김수영(2006)의 69.6-80%, Kim(2011)의 72.6-96%, 허양희·설훈구·홍재범(2010)의 75.7-80.1% 등의 기존의 단독 분류방법에 의한 연구결과를 월등하게 능가하는 것이다.

일반적으로 예측모형의 정확성을 평가할 때 전체모형의 예측을 만을 고려하는 경향이 있는데, 도산기업을 비도산기업으로 분류하는 type I 오류율이 더 많은 비

용을 초래한다는 점에 근거하여(Ananarajan *et al.*, 2001; Sinkey, 1975), type I 오류율을 눈여겨 볼 필요가 있다. 본 연구에서 제시한 SVM-NN hybrid 모형은 type I 오류율이 0%로 전체예측률 뿐 아니라 Type I 오류율 측면에서도 예측정확성이 비약적으로 향상되었다.

〈표 3〉 모형의 예측정확성: SVM-NN hybrid 모형

		도산상태	예측 상황		총계
			비도산	도산	
실제 상황	표본수	비도산	76	6	82
		도산	0	82	82
분류정확율(%)		비도산	92.68	7.32	100
		도산	0	100	100

주: 모형의 전체예측율 = 96.34 %

모형 내 입력노드의 상대적 중요성을 나타내는 상대치를 〈표 4〉에 제시하였다. SVM-NN 분석모형은 매출채권회전율과 총자본경상이익율, 당좌비율, 고정장기적합율, 부채비율, 고정자산회전율, 매출액순이익율, 총자산증가율의 여덟 개의 예측변수를 포함한다. 수익성비율에서는 총자본경상이익율과 매출액순이익율이 유의적인 변수로 파악되었으며, Falbo(1991)와 Keating *et al.*(2005), Song(1987), Tuckman & Chang(1991)의 연구결과와 동일하게 수익성과 도산가능성 간에 부의 관계가 존재함을 확인하였다. 특히 SVM-NN hybrid모형에 의해 그 상대적 중요성이 부각된 예측변수로는 총자본경상이익율이 있는데, 이는 기업에 투자 운영된 총자본의 수익성이 호텔의 도산여부에 유의적인 영향을 미친다는 것을 의미한다. 총자본경상이익율과 매출액순이익율, 자기자본경상이익율, 고정자산회전율을 함께 고려해 볼 때, 도산호텔의 낮은 수익성은 저조한 매출과 마케팅 및 구매기능에 대한 비효율적인 통제에서 야기될 수 있음을 확인하였다.

〈표 4〉 SVM-NN 모형 내 입력노드의 상대적 중요성

	입력노드	상대적 중요성
수익성비율	총자본경상이익율	0.00885
	매출액순이익율	0.00236
유동성비율	당좌비율	0.00688
	매출채권회전율	0.01990
활동성비율	고정자산회전율	0.00360
안정성비율	부채비율	0.00405
	고정장기적합율	0.00471
성장성비율	총자산증가율	0.00063

유동성비율에서는 당좌비율과 매출채권회전율이 부도에 유의적인 영향을 미치는 것으로 확인하였다. 매출채권회전율은 Newton(2010)과 Tamari(1964)의 연구결과와 같이 부도가능성과 정의 관계를 갖고 있으며, 반면에 당좌비율은 Keasey & McGuinness(1990)와 Keasey, McGuinness, & Short(1990), Lee *et al.*(1996), Peel & Peel(1987), Zheng & Yanjun(2010)의 연구결과와 유사하게 부의 관계를 보이고 있다. 활동성비율 범주의 고정자산회전율을 안정성비율의 부채비율과 고정장기적합율과 더불어 분석하면, 도산호텔들이 상대적으로 부채에 대한 의존성이 높은 것을 알 수 있다. 자본구조상의 높은 부채비율은 특히 경제 불황기 동안에는 레버지지를 최대화하기보다는 금융구조를 악화시켜 호텔의 도산가능성을 높이는 것으로 판단된다. 마지막으로, Doumpos, Gaganis & Pasiouras(2005)와 Gaganis, Pasiouras, Spathis & Zopounidis(2007)의 연구결과와 유사하게 부도호텔의 총자산증가율이 비도산 호텔에 비해 낮은 것으로 파악되었다. 이를 상대적으로 낮은 부도호텔의 자기자본증가율과 복합적으로 고려해 볼 때, 부도호텔들이 매출채권이나 장기투자자산 등을 처분하여 레버리지에 대한 구조조정의 노력을 피하고 있으나 호텔도산의 본질적인 문제를 해결하지 않고는 도움이 되지 않는다는 것을 알 수 있다.

McNemar 테스트를 통해 본 연구에서 제시한 SVM-NN hybrid모형이 기타의 단일 모형의 예측정확성을 능가하는가를 확인하였으며 그 결과를 <표 5>에 제시하였다. 이는 Chi-square 분포를 활용하여 두 연관된 그룹에 대해 비모수적(non-parametric) 테스트를 하는 것(Hollander & Wolfe, 1999; Krishnamoorthy, Thomson, & Cai, 2004)으로 독립적인 두 집단 간의 유의적 상이성을 평가한다. <표 5>에서 볼 수 있듯이 SVM-NN Hybrid 모형은 SVM, NN의 독립적인 예측방법보다 우수한 예측력을 보여주고 있다.

<표 5> 분류방법 간의 McNemar 테스트

	NN	SVM	SVM-NN	
NN		0.1020	0.00154	**
SVM			0.02118	**

주: *** p<0.01: ** p<0.05: *p<0.10

3. 모형의 실질적 응용

호텔 채권자나 소유주는 본 연구에서 제시하는 예측모형을 이용하여 기존의 운

영호텔이나 새롭게 인수하고자 하는 호텔의 도산을 예측할 수 있다. 나아가 도산 호텔의 미결제 부채에 대한 채무를 근거로 투자하고자 하는 호텔의 부채와 자산물건의 순가에 의한 시장가치를 평가할 수 있다. 경영진 역시 모형의 결과를 활용하여 호텔이 직면하게 될 도산가능성을 확인하고 이에 대해 다양한 자구책을 사전에 마련할 수 있다.

도산과 경제의 구조조정에 대한 중요성이 부각되면서 기업도산과 관련한 이론과 경험, 규범적인 것을 총괄하는 학술연구들이 많이 진행되어왔다. 미국의 많은 대학들은 컨퍼런스와 논문발표를 지원해왔으며, 1992년 Financial Management Association에서는 재무부문에 도산과 부실기업에 대한 별도의 토론세션을 제공하였다. 국내 역시 기업도산이 미치는 경제, 사회적 영향력을 감안하여 관련 컨퍼런스에서 기업도산의 법적인 측면과 재무적인 측면을 논의할 수 있다. 나아가 국내호텔도산에 대한 연구는 연구논문 및 학술저서, 석·박사 학위논문, 기타 공식적인 분석을 위한 주요 주제가 될 수 있다. 특히 최근의 경제 불황기에 호텔기업의 도산율이 증가되는 것을 감안할 때, 호텔재무회계 강의에 본 주제를 포함시킴으로써 미래 호텔기업의 주역들에게 호텔도산이 기업과 사회에 미칠 수 있는 영향을 충분히 인지하고 이를 방지하기 위한 경영정책을 수립할 수 있는 능력을 배양할 수 있다.

V. 결론 및 미래연구에 대한 제언

도산은 높은 비용을 야기하는 경제 사회적 문제이다. O'Leary (1998)가 도산 예측을 사업 의사결정 문제 중 가장 중요한 이슈 중의 하나로 언급하였음에도 불구하고, 가장 정확한 도산예측의 기술에 대해서는 결론이 나지 않은 상태이다. 본 연구는 SVM-NN hybrid모형을 사용한 호텔도산예측모형을 개발하는데 목적을 두었다. 국내 관광호텔의 자료를 사용한 실험적인 연구 결과, SVM-NN hybrid모형이 특히 경제 불황기에는 도산의 위험이 있는 호텔들에 대해 보다 예민한 예측변수를 제공할 뿐 만 아니라, 전체 예측율을 96%이상으로 개선하고, type I 오류율을 제거하여 비용효율적인 예측 모형을 제공하고 있음을 증명하였다. 이는 도산예측의 정확성이 약간만 개선되어도 호텔투자자나 소유자의 수익에 큰 영향을 미칠 수 있다는 점에서 중요한 의미를 갖는다.

본 연구가 갖는 한계로는 우선 특정 9년 기간 동안의 국내 관광호텔을 표본으로

연구하였으므로, 다른 국가나 다른 시점에 대해 연구결과를 일반화하는 것에 제한적일 수 있다. 또한 본 연구는 재무요소 만을 고려하여 호텔도산예측모형을 구축하였으므로, 호텔 연한이나 시장가치 등의 기타의 계량적인 요소와 경영능력이나 소유권형태 등의 정성적인 요인을 포함한 추가연구가 가능하다. 여러 연구를 통해 확인되었듯이 현재 사용하는 여러 가지 방법들을 지속적으로 능가하는 새로운 접근방법을 개발한다는 것은 쉬운 일이 아니다. 예측이 보다 정확하게 이루어지는 것은 복합적인 학습접근방법이 사용하는 자료의 특성에 근거하여 어느 정도까지 예민하게 적용될 수 있는지에 의존하기 때문이다. 따라서 경제상황이나 제반 리스크 요인 등, 호텔도산과 밀접한 상관관계를 가지고 있는 기타 변수를 포함하여 최대한 많은 수의 호텔데이터와 긴 평가기간을 포함함으로써 특이치를 흡수하여 통계적 의미를 갖도록 노력하여야 할 것이다. 그 외에도 다양한 데이터 유형과 구조를 지닌 ensemble모형구축 등을 미래연구에 포함할 수 있다. 마지막으로 잠재적 문제점을 사전에 예방한다는 측면에서, 호텔산업 뿐 아니라 외식산업 및 컨벤션산업 등의 연관 산업에 대해서도 도산예측모형을 개발하여 널리 활용할 수 있기를 기대해 본다.

참고문헌

- 김수영(2006). 다변량 판별분석과 로지스틱 회귀분석. 인공신경망 분석을 이용한 호텔도산 예측. 『관광학연구』, 30(2), 53-85.
- 김수정·김우곤(1998). 호텔산업의 도산원인 분석. 『관광학연구』, 22(2), 102-109.
- 김시중(2005). 호텔기업 부실예측모형의 예측력 비교연구. 『관광학연구』, 28(4), 9-26.
- 통계청(2011). 2004~2009 사업체 생성·소멸(생멸) 현황 분석. http://kostat.go.kr/portal/korea/kor_nw/2/1/index.board?bmode=read&aSeq=245574
- 한국은행경제통계시스템. <http://ecos.bok.or.kr/>. 2011.09.01.
- 허양희·설훈구·홍재범. (2010). 관광호텔산업에서의 부실예측모형개발에 관한 실증연구. 『관광·레저연구』, 22(6), 253-270.
- Abbss, H., Towsey, M., & Finn, G.(2001). C-Net: A method for generating non-deterministic and dynamic multivariate decision trees. *Knowledge and Information Systems* 3(2), 184-197.
- Assaad, M., Bone, R., & Cardot, H.(2008). A new boosting algorithm for improved time-series forecasting with recurrent neural networks.

Information Fusion, 9, 41-55.

- Aziz, M., & Dar, H.(2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance*, 6(1),18-33.
- Beaver, W.(1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 5(Suppl), 71-102.
- Booth, P.(1983). Decomposition measures and the prediction of financial failure. *Journal of Business Finance and Accounting*, 10(1), 67-82.
- Boritz, J., Kennedy, D. & Albuquerque, A.(1995). Predicting corporate failure suing a neural network approach. *Intelligent Systems in Accounting, Finance and Management*, 4(2), 95-111.
- Braun, H., & Chandler, J.(1987). Predicting stock market behavior through rule induction: An application of the learning-from-example approach. *Decision Sciences*, 18(3), 415 - 429.
- Breiman, L.(1996a). Bagging predictors, *Machine Learning*, 24(2), 123-140.
- _____ (1996b). *Stacked regressions*. *Machine Learning*, 24(1), 49-64.
- Charlambous, C., Charitou, A., & Kaourou, F.(2000). Comparative analysis of artificial neural network models: Application in bankruptcy prediction. *Annals of Operational Research*, 99(4), 403-425.
- Chen, W., Ma, C., & Ma, L.(2009). Mining the customer credit using hybrid support vector machine technique. *Expert Systems with Applications*, 36(4), 7611-7616.
- Cho, S., Ahn, J., & Lee, S.(2001). Exploiting diversity of neural networks with special evolution. *Proceedings of the international joint conference on neural network (IJCNN'01)*. Washington DC, July, 808-813.
- Cho, S., Kim, J., & Bae, J.(2009). An integrative model with subject weight based on neural network learning for bankruptcy prediction. *Expert Systems with Applications*, 36(1), 403-410.
- Cristiamini, N., & Shawe-Taylor, J.(2000). *An introduction to support vector machines*. Cambridge, UK: Cambridge University Press.
- Cybenko, G.(1989). Approximation by Superpositions of a Sigmoidal Function. *Mathematical Control Signal Systems*, 2(4), 303-314.
- Davalos, S., Gritta, R. & Chow, G.(1999). The application of a neural network approach to predicting bankruptcy risks facing the major US air carriers: 1979-1996. *Journal of Air Transport Management*, 5(1), 81-86.

- Deakin, E. B.(1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research*, 10(1), 167-179.
- Dietterich, T.(2000). Experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. *Machine Learning*, 40(2), 139-157.
- Dimitras, A., Zanakis, S., & Zopounidis, C.(1996). A survey of business failure with an emphasis on prediction methods and industrial application. *European Journal of Operational Research*, 90(3), 487-513.
- Doumpos, M., Gaganis, C., & Pasiouras, F.(2005). Explaining qualifications in audit reports using a support vector machine methodology. *Intelligent Systems in Accounting, Finance and Management*, 13(4), 197-215.
- Doran, M., Raicu, D., Furst, J., Settimi, R., Schipma, M., & Chandler, D.(2007). Oligonucleotide micro array identification of bacillus anthracis strains using support vector machines. *Bio informatics*, 23(4), 487 - 492.
- Duin, R., & Tax, D.(2000). Experiments with classifier combination rules, In Kittler, J., & Roli, F., editors, Multiple Classifier Systems. First International Workshop, MCS, Cagliari, Italy, 1857 of *Lecture Notes in Computer Science*. Springer-Verlag, 16-29.
- Dun and Bradstreet Corporation.(1994). *Business Failure Record*. Author, New York.
- Efon, B., & Gong, G.(1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37(1), 36-48.
- Esteban, A., Garcia, N., Gamez, M., & Elizondo, D.(2007). Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decision Support Systems*, 45(1), 110-122.
- Etheridge, H., & Sriram, R.(1997). A comparison of the relative costs of financial distress models: Artificial neural networks, logit and multivariate discriminant analysis. *Intelligent Systems in Accounting, Finance and Management* 6(3), 235-248.
- Falbo, P.(1991). Credit scoring by enlarged discriminant analysis. *OMEGA*, 19(4), 275-289.
- Fletcher, D., & Goss, E.(1993). Forecasting with neural networks: An application using bankruptcy data. *Information and Management*,

24(3), 159-167.

- Freund, Y., & Schapire, R.(1996). Experiments with a new boosting algorithm. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence* Detroit, MI., 788-793.
- Gaganis, C., Pasiouras, F., Spathis, C., & Zopounidis, C.(2007). A comparison of nearest neighbors, discriminant and logit models for auditing decisions. *Intelligent Systems in Accounting, Finance and Management* 15(1), 23-40.
- Gao, L.(1999). Study of business failure in the hospitality industry from both microeconomic and macroeconomic perspectives. Master thesis in hotel administration, University of Nevada, Las Vegas, USA.
- Grice, J., & Dugan, M. (2001). The limitations of bankruptcy prediction models: Some cautions for the researcher. *Review of Quantitative Finance and Accounting*, 17(2), 151-166.
- Goh, A.(1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143-151.
- Gu, Z.(2002). Analyzing bankruptcy in the restaurant industry: A multiple discriminant model. *International Journal of Hospitality Management*, 21(1), 25 - 42.
- Hambrick, D., & D'Aveni, R.(1988). Large corporate failures as downward spirals. *Administrative Science Quarterly*, 33(1), 1-23.
- Hammer, M.(1983). Failure prediction: sensitivity of classification accuracy to alternative statical methods and variable sets. *Journal of Accounting and Public Policy*, 2(4), 289-307.
- Hanson, R.(2003). A study of Altman's revised four-variable Z-score bankruptcy prediction model as it applies to the service industry. PhD dissertation in business and entrepreneurship, Nova Southeastern University, USA.
- Hastie, T., Tibshirani, R., Friedman, J.(2001). *The Elements of Statistical Learning-Data mining, Inference and Prediction*. Springer series in Statistics.
- Hollander, M., & Wolfe, D.A.(1999). Comparing two success probabilities. In M. Hollander & D.A Wolfe (Eds.), *Nonparametric statistical methods* (2nd ed.) New York: Wiley-Interscience, pp458 - 494.
- Hornik, K., Stinchcombe, M., & White, H.(1989). Multilayer feed forward

networks are universal approximations. *Neural Networks*, 2(5), 336-359.

- Hua, Z., Wang, Y., Xu, X., Zhang, B., & Liang, L.(2007). Prediction corporate financial distress based on integration of support vector machine and logistic regression. *Expert Systems with Applications*, 33(2), 434-440.
- Huang, Z., Chen, H., Hsu, C., Chen, W., & Wu, S.(2004). Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision Support Systems*, 37(4),543-558.
- Hung, C., & Chen, J.(2009). A selective ensemble based on expected probabilities for bankruptcy prediction. *Expert Systems with Applications*, 36(3), 5297-5303.
- Jhee, W., & Lee, J.(1993). Performance of neural networks in managerial forecasting. *Intelligent systems in Accounting, Finance, and Management*, 2(1), 55 -71.
- Jo, H., & Han, I.(1996). Integration of case-based forecasting, neural network, and discriminant analysis for bankruptcy prediction. *Expert Systems with Applications*, 11(4), 415-422.
- Jo, H., Han, I., & Lee, H.(1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems with Applications*, 13(2), 97-108.
- Keasey, K., & McGuinness, P.(1990). The failure of UK industrial firms for the period 1976-1984: Logistic analysis and entropy measures. *Journal of Business Finance and Accounting*, 17(1), 119-135.
- Keasey, K., McGuinness, P., & Short, H.(1990). Multilogit approach to predicting corporate failure- further analysis and the issue of signal consistency, *OMEGA*, 18(1), 85-94.
- Keating, E., Fischer, M., Gordon, T., & Greenlee, J.(2005). Assessing financial vulnerability in the nonprofit sector. KSG working paper No. RWP05-002: Hauser center for nonprofit organizations Paper No. 27. < http://papers.ssrn.com/sol3/papers.cfm?abstract_id=647662>.
- Kim, H., & Gu, Z.(2006). Predicting restaurant bankruptcy: A logit model in comparison with a discriminant model. *Journal of Hospitality and Tourism Research*, 30(4), 474-493.
- Kim, K.(2003). Financial time series forecasting using support vector

machines. *Neuro Computing*, 55(1/2), 307-319.

- Kim, S.(2011). Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. *The Service Industries Journal*, 31(3-4), 441-468.
- Kittler, J.(2000). A framework for classifier fusion: Is it still needed? *LNCS*, 1876, 45-56.
- _____ (1998). Combining classifiers: A theoretical framework. *Pattern Analysis and Application*, 1(1), 18-27.
- Koh, H.(2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal*, 19(3), 462-476.
- Krishnamoorthy, K., Thomson, J., & Cai, Y.(2004). An exact method of testing equality of several binomial proportions to a specified standard. *Computational Statistics and Data Analysis*, 45(4), 697-707.
- Lahiri, R.(2006). Comparison of data mining and statistical techniques for classification model. PhD dissertation in information systems and decision science, Jadavpur University, India.
- Ledezma, A., Aler, R., Sanchis, A., & Borrajo, D.(2010). GA-Stacking: evolutionary stacked generalization. *Intelligent Data Analysis*, 14(1), 89-119.
- Lee, K., Han, I., & Kwon, Y.(1996). Hybrid neural network models for bankruptcy prediction. *Decision Support Systems*, 18(1), 63-72.
- Lee, S., & Oh, S.(1990). A comparative study of recursive partitioning algorithm and analog concept learning systems. *Expert Systems with Applications*, 1(4), 403-416.
- Lennox, C.(1999). Identifying failing companies: A reservation of the logit, probit and DA approaches. *Journal of Economics and Business*, 51(4), 347-364.
- Leshno, M., & Spector, Y.(1996). Neural network prediction analysis: The bankruptcy case. *Neuro computing*, 10(2), 125-147.
- Lev, B.(1969). *Accounting and information theory*. Evanston, Illinois: American Accounting Association.
- Li, M., Sun, J., & Sun, B.(2009). Financial distress prediction based on OR-CBR in the principle of k-nearest neighbors. *Expert Systems with Applications*, 36(1), 643-659.
- Lin, F., & McClean, S.(2001). A data mining approach to the prediction of

corporate failure. *Knowledge-Based Systems*, 14(3-4), 189-195.

- Lipovatz, D., Mandaaraka, M., & Mourelatos, A.(2000). Multivariate analysis for the assessment of factors affecting industrial competitiveness: The case of Greek food and beverage industries. *Applied Stochastic Models in Business and Industry*, 16(2), 85-98.
- Malhotra, R., & Jain, A.(2010). Software effort prediction using statistical and machine learning methods. *International Journal of Advanced Computer Science and Applications*, 2(1), 145-152.
- Maqsood, I., Khan, M., & Abraham, A.(2004). An ensemble of neural networks for weather forecasting. *Neural Comput and Application*, 13(2), 112-122.
- McKee, T. E., & Lensberg, T.(2002). Genetic programming and rough sets: A hybrid approach to bankruptcy classification. *European Journal of Operational Research*, 138(2), 436-451.
- Mei, A., Shen, Q., Ye, B.(2008). Hybridized KNN and SVM for gene expression data classification. *Life Science Journal*, 6(1), 61-66.
- Min, J., & Lee, Y.(2005). Bankruptcy prediction using support vector machine with optimal choice of kernel-function parameters. *Expert Systems with Applications*, 28(4), 603-614.
- Min, S., Lee, J., & Han, I.(2006). Hybrid genetic algorithms and support vector machines for bankruptcy prediction. *Expert Systems with Applications*, 31(3), 652-660.
- Moncarz, E., & Kron, R.(1993). Operational analysis: A case study of two hotels in financial distress. *International Journal of Hospitality Management*, 12(2), 175-196.
- Mukkamala, S., Janowski, G., & Sung, A.(2001). Intrusion detection neural networks and support vector machines. Hybrid Information Systems Advances in *Soft Computing*, Heidelberg, Germany. Springer, 121-138.
- Nanda, S., & Pendharkar, P.(2001). Linear models for minimizing misclassification costs in bankruptcy prediction. *International Journal of Intelligent Systems in Accounting, Finance, and Management*, 10(3), 155-168.
- Newton, G.(2010). Economic causes of business failure. In *Bankruptcy and Insolvency Accounting* (pp.49-51). New Jersey: John Wiley and Sons, Inc.

- Noronha, G., & Singal, V.(2004). Financial health and airline safety. *Managerial and Decision Economics*, 25(1), 1-16.
- Odom, M., & Sharda, R.(1993). A neural network model for bankruptcy prediction. In Trippi, R. & Turban, E. (Eds.), *Neural networks in finance and investing*. Chicago-London: Probus, pp.177-185.
- Ohlson, J.(1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109-131.
- O'Leary, D.(1998). Using neural networks to predict corporate failure. *International Journal of Intelligent Systems in Accounting, Finance, and Management* 7(3), 187-197.
- Olmeda, I., & Fernandez, E.(1997). Hybrid classifiers for financial multicriteria decision making: The case of bankruptcy prediction. *Computational Economics*, 10(4), 317 - 335.
- Optiz, D., & Maclin, R.(1999). Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, 11, 169-198.
- Pal, M., & Mather, M.(2004). Assessment of the effectiveness of support vector machines for hyper spectral data. *Future Generation Computer Systems*, 20(7), 1215 - 1225.
- Parsa, H., Self, J., Njite, D., & King, T.(2005). Why restaurant fail. *Cornell Hotel and Restaurant Administration Quarterly*, 46(3), 304 - 322.
- Patterson, D.(1971). Bankruptcy prediction in the casino industry. Master thesis in hotel administration, University of Nevada, Las Vegas, USA.
- Peel, M., & Peel, D.(1987). Some further empirical evidence on predicting private company failure. *Accounting and Business Research*, 18/69, 57-66.
- Pendharkar, P.(2001). An empirical study of designing and testing of hybrid evolutionary-neural approach for classification. *Omega*, 29(4), 361-374.
- Raghupathi, W., Schkade, L., & Raju, B.(1991). A neural network approach to bankruptcy prediction, *Proceedings of the IEEE 24th Annual Hawaii International Conference on Systems Sciences*, 4, 147-155.
- Ramaswamy, P., Tamayo, R., Rifkin, S., Mukherjee, C., Yeang, M., Angelo, C., Ladd, M., Reich, E., Latulippe, J., Mesirov, T., Poggio, W., Gerald, M., Loda, E., Lander, T., & Golub.(2001). Multiclass cancer diagnosis using tumor gene expression signatures. *Proceedings of the*

National Academy of Sciences of the United States of America. 98(26), 15149-54.

- Ravi Kumar, R., & Ravi, V.(2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques—a review. *European Journal of Operational Research*, 180(1), 1 - 8.
- Rong, L. & Yuan, S.(2010), Diagnosis of breast tumor using SVM-KNN classifier. *2010 Second WRI Congress on Intelligent Systems*, 3, 95-97.
- Rumelhart, D., Hinton, G., & Williams, R.(1986). Learning internal representations by error propagation. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel distributed processing* (Vol. 1). Cambridge, MA: MIT Press.
- Schapire, R.(1990). The strength of weak learnability. *Machine Learning*, 5(2), 197-227.
- Schmitz, G., Aldrich, C., Gouws, F.(1999). ANN-DT: An algorithm for extraction of decision trees from artificial neural networks. *IEEE Transactions of Neural Networks*, 10(6), 1392-1401.
- Shin, K., Lee, T., & Kim, H.(2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1), 127 - 135.
- Shachmurove, Y.(2002). Applying artificial neural networks to business, economics and finance (Penn CARESS Working Papers). Los Angeles, CA: UCLA Department of Economics.
- Surkan, A., & Singleton, C.(1990). Neural networks for bond ratings improved by multiple hidden layers. *Proceedings of IEEE International Conference on Neural Networks*, San Diego, USA., 162 - 168.
- Song, I.(1987). An empirical study on relation of business failure reasons through reformulation of bankruptcy prediction model. Sungkyunkwan University 113-142.
- Sung, T., Chang, N., & Lee, G.(1999). Dynamics of Modeling in data mining: Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, 16(1), 63-85.
- Swicegood, P., & Clark, J.(2001). Off-site monitoring systems for predicting bank under performance: A comparison of neural networks, discriminant analysis, and professional human judgment, *International Journal of*

Intelligent Systems in Accounting, Finance & Management, 10(3), 169-186.

- Tam K., & Kiang, M.(1992). Managerial applications of neural networks: The case of bank failure predictions. *Management Science*, 38(7), 926-947.
- Tamari, M.(1964). *Financial ratios as a means of forecasting bankruptcy*, *Economic Review*, Bank of Israel, Jerusalem.
- Tavlin, E. M., Moncrz, E., & Dumont, D.(1989) Financial failure in the hospitality industry. *FIU Review*, 7(1), 55-75.
- Tsukuda, J., & Baba, S.(1994). Prediction Japanese corporate bankruptcy in terms of financial data using neural network. *Computers and Industrial Engineering*, 27(1-4), 445-448.
- Tuckman, H. & Chang, C.(1991). A methodology for measuring the financial vulnerability of charitable nonprofit organization. *Nonprofit and Voluntary Sector Quarterly*, 20(4), 445-460.
- Valentini, G., & Masulli, F.(2002). Ensembles of learning machines. Neural Nets WIRNVietri-02, *Series Lecture Notes in Computer Science* 2486, Springer-Verlag, Heidelberg, Germany, 3-19.
- Ulas, M.(2009). Incremental construction of cost-conscious ensembles using multiple learners and representations in machine learning, PhD. Dissertation for computer engineering, Bogazici University, Istanbul, Turkey.
- Vapnik, N.(1995). *The nature of statistical learning theory*. New York: Springer.
- _____ (1998). *Statistical learning theory*. New York: Wiley.
- Verwoerd, T., & Hunt, R.(2002). Intrusion detection techniques and approaches. *Computer Communication*, 25(15), 1356-1365.
- Wang, H., Zhao, T., Tan, H., & Zhang, S.(2008). Biomedical named entity recognition based on classifiers ensemble. *International Journal of Computer Science and Applications*, 5(2), 1-11.
- West, D., Dellana, S., & Qian, J.(2005). Neural network ensemble strategies for financial decision applications. *Computers and Operations Research*, 32(10), 2543-2559.
- Wolpert, D. H.(1992). Stacked generalization. *Neural Networks*, 5, 241-259.
- Wong, B., Bondnovich, T., & Yakup, S.(1997). Neural network applications

- in business: A review and analysis of the literature (1988 - 95). *Decision Support Systems*, 19(4), 301 - 320.
- Wu, C., Tzeng, G., Goo, Y., & Fang, W.(2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert Systems with Applications*, 32(2), 397 - 408.
- Yang, S., & Browne, A.(2004). Neural network ensembles: combining multiple models for enhanced performance using a multi state approach. *Expert Systems*, 21(5), 279-288.
- Yim, J., & Mitchell, H.(2002). A comparison of Australian financial service failure models: Hybrid neural networks, logit models and discriminant analysis. *RMIT working paper*, ISSN 1038-7448, 1-35.
- Yoon, Y., Guimaraes, T., & Swales, G.(1994). Integrating artificial neural networks with rule-based expert systems. *Decision Support Systems*, 11(5), 497-507.
- Youn, H., & Gu, Z.(2011). Predict US restaurant firm failures: The artificial neural network model versus logistic regression model, *Tourism and Hospitality Research*, 10(3). 171-187.
- Zheng, Q., Yanjun, C.(2010). The Bankruptcy Prediction of Chinese Export-oriented Enterprise: base on the Financial Crisis. *International Journal of Trade ,Economics and Finance*, 1(3), 283-288.
- Zhang, H., Berg, A., Maire, M., & Malik, J.(2006), SVM-KNN: Discriminative nearest neighbor classification for visual category. *IEEE Conference on Computer Vision and Pattern Recognition*, 2, 2126-2134.
- Zhou, J., & Lopresti, D.(1997). Extracting Text from WWW Images. *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, August, Ulm, Germany, 248-252.
- Zmijewski, M.(1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22(suppl.), 59-86.