

# DEA 와 SVM 을 통합한 IT 벤처기업의 효율성 평가

홍태호<sup>a</sup>, 박지영<sup>b</sup>

<sup>a</sup> 부산대학교 상과대학 경영학부  
609-735, 부산광역시 금정구 장전동 산 30  
Tel: +82-51-510-1361, Fax: +82-51-581-3142, E-mail: hongth@pusan.ac.kr

<sup>b</sup> 부산대학교 대학원 경영학과  
609-735, 부산광역시 금정구 장전동 산 30  
Tel: +82-51-510-2531, E-mail: jiyoun@pusan.ac.kr

## Abstract

IT 벤처기업은 자본 대비 높은 수익성을 가지므로 지식기반 산업환경에서 많은 투자자들의 집중적인 관심을 받고 있다. 이러한 IT 벤처기업의 효율성을 평가하기 위한 방안으로, DEA 와 데이터마이닝 기법을 통합하는 방안을 제시하였다. 국내 코스닥 상장기업 가운데 IT 에 주력하고 있는 벤처기업들을 대상으로 본 연구에서 제시한 효율성 평가방법을 적용하였다. 대표적인 비모수적 분석기법인 Data Envelopment Analysis(DEA)를 이용하여 연구대상 기업들을 효율기업 및 비효율기업으로 구분한 후, DEA 의 효율성을 설명하는 모형을 logit 을 이용하여 구축하였다. DEA 는 기업의 상대적인 효율성을 측정하는 데에는 우수하지만, 효율성 정도를 설명하는 모형의 구축에는 한계가 있다. 이를 보완한 DEA 의 결과를 logit 과 통합한 효율성 모형에 대해서 데이터 마이닝 기법인 logit, 판별분석, Support Vector Machine(SVM) 등을 적용하여 IT 벤처기업의 효율성을 사전에 예측하여 평가 및 투자에 활용할 수 있는 방안을 제시하였다.

## Keywords:

IT venture business; Efficiency; DEA; SVM

## 서론

IT 벤처기업들은 그들의 가장 큰 자산인 정보기술을 통해 수익 면에서 상당한 이점을 누리고 있지만, 벤처기업의 환경적 특성으로 인해 경영상태의 불확실성도 크다고 할 수 있다. 특히, 지난 2000 년 하반기에 닷컴기업과 벤처버블의 붕괴위기와 더불어 국내 소프트웨어 기업들은 어려운 경제상황에 처했다(김우식 외, 2003). 벤처기업들은 금융위기 이후에 양적으로 매우 빠른 성장을 하였고, 정부차원의 지원을 통해 사회적으로도 많은 영향을 주었다. 2000 년 초반에 비하여 현재 벤처기업의 수가 현저하게 감소

하는 이유는, 벤처기업의 요건이 지속적으로 강화됨으로써, 수많은 부실 벤처기업이 퇴출되었기 때문으로 분석된다(조준희, 2003). 벤처기업이 처한 환경적 특성 중 하나는 경영환경의 불확실성이라 할 수 있다. 때문에 많은 벤처기업들은 자사의 기술 및 자산 등에 대한 효율적 운영 문제 등 경영성과와 관련된 사항에 대한 다각적인 논의가 필요하며, 이에 따른 체계적인 전략수립도 필요하다. 투자자들의 입장에서는 어떤 벤처기업이 경쟁적으로 이점이 있는지를 알고 싶어하며, 투자자들은 수익을 얻을 수 있는 기업에 투자하려고 한다. 이러한 관점에서, 동종업계에 관심기업이 효율적으로 운영되고 있는가 그렇지 않은가에 대해 알 수 있다면, 매우 유익한 정보가 될 수 있을 것이다.

본 연구에서는 이러한 연구문제를 바탕으로 하여 코스닥에 상장된 소프트웨어 서비스 관련 IT 벤처기업들의 2005 년도 재무정보를 이용하여 이들의 상대적 효율성 평가를 통해 효율성이 높은 기업과 그렇지 않은 기업들의 특성을 비교해보고자 한다. 또한 효율적인 기업들에게 상대적으로 중요도가 높은 재무정보가 무엇인지 도출해내고자 한다. 대상 기업들의 상대적 효율성 평가를 위해, 대표적인 비모수적 접근법인 Data Envelopment Analysis(DEA)를 이용하고, DEA 를 통해 상대적으로 효율적이라고 평가된 기업들을 특징짓는 재무변수가 무엇인지 로짓분석을 통해 알아보고자 한다. 이러한 DEA-logit 통합모형을 이용하면 다른 기업들과의 상대적 비교 없이 특정 IT 벤처기업의 효율성 평가가 가능하게 된다. 또한, 효율기업예측을 위해 우수한 기계학습기법으로 주목받고 있는 Support Vector Machine(SVM) 모형을 적용하고자 한다. 이상과 같이 연구문제를 정리하면 다음과 같이 두 가지로 요약될 수 있다:

1) DEA 의 효율성을 설명하는 모형이 가능한가?

DEA 는 상대적인 효율성을 측정하는 데 우수한 능력을 갖고 있다. 하지만, 벤처기업의 투자자 또는 금융기업의 평가자는 특정 기업의 효율성을 평가할 수 있는 모형이 필요하다. 이와 같이, DEA 의 효율성을 설명하는 모형의 개발을 할 필요가 있으면 그 대안으로 로짓모형이 적합한지를 밝히고자 한다.

2) 벤처기업 평가 및 투자에 활용할 수 있는 예측모형의 개발이 가능한가?

DEA의 사후적으로 상대적 효율성을 측정할 수 있는 방법이다. 벤처기업 투자자 또는 금융기관의 평가자에게 의사결정에 활용할 수 벤처기업의 효율성 예측모형을 개발하고자 한다.

본 연구의 구성은 다음과 같다. II장에서는 IT 벤처기업의 특성과 함께, 기업의 상대적 효율성 측정을 위해 사용되는 DEA 모형 및 관련 연구와 SVM에 대해 설명한다. III장에서는 본 연구의 연구모형을 제시하고, IV장에서는 DEA 및 로짓분석, SVM 등의 실증분석 결과를 제시하며, 마지막으로 V장에서는 연구결과 및 향후 연구 과제에 대해 논의하기로 한다.

## 선행연구

### IT 벤처기업의 특성

제조기반 산업에서 지식기반 산업으로의 급속한 변화는 핵심지식을 바탕으로 하는 벤처기업에 대한 관심을 더욱 더 고조시키는 계기가 되었으며, 이러한 관심이 고조되면서 벤처기업의 성공요인에 대한 분석 및 개념적 특성과 본질에 관한 연구 등도 지속되고 있다. 특히, 벤처기업 가운데 High Tech, High Risk, High Return을 추구하는 첨단 벤처기업의 대부분은 정보통신분야에 있어서도 특히 소프트웨어 부문에 집중되어 있다(김우식 외, 2003). 본 연구의 대상인 IT 벤처기업은 특히 창조적인 아이디어를 활용할 수 있는 기술집약적 분야에서 더욱 활발한 활동을 하고 있다. 제조기반 산업에서 보다, 자본 대비 높은 수익성을 가지기 때문에, 지식기반 산업 환경에서 그 매력도는 상당하다고 할 수 있지만, 새로운 기술과 아이디어 측면이 주요 기반이 되므로 수요나 기술 측면에 있어 위험이 따를 수 있기 때문에, 벤처기업에 대한 정확한 특성을 파악하여 위험은 최소화하고, 상대적으로 이익을 더 높일 수 있는 것들에 대해서는 전략적으로 개선해나가야 할 필요가 있을 것이다.

IT 벤처기업의 효율성 분석에 앞서, 벤처기업에 대한 정의 및 특성에 관한 분류 범위의 한계는 어디까지인지에 대한 고찰이 필요하다. 벤처기업의 정의 및 특성은 매우 광범위하며 다양하게 사용되고 있다. 이진주(1988)는 벤처기업을 새로운 아이디어와 기술력으로 사업에 도전하는 모험적인 중소기업으로 정의하면서, 소수의 기술창업자가 기술혁신의 아이디어를 상업화하기 위해서 설립한 신생기업의 경우가 많기 때문에 위험부담이 높을 수 있으나, 성공할 경우에는 높은 기대이익이 발생된다는 특징을 지적하였다. 이장우(1998)는 새로운 아이디어와 기술을 가지고 사업에 도전하는 모험적인 중소기업을 벤처기업으로 지칭하였으며, 고위험과 고성과를 특징으로

하는 기술집약적인 중소기업으로 정의하였다. 전체적인 특성으로 정의하면, 벤처기업은 포괄적인 정의로 모험사업이며, 구체적인 정의로 신생기업이라는 점을 강조할 수 있다(서종상 외, 2006). 이준우(2005)는 연구대상으로서 벤처기업의 정의와 범위에 대한 학술적인 연구 노력이 매우 미흡함을 지적하고 미국, 일본, 한국의 학술 연구 및 실무 정책에 있어 강조 측면에 따라 벤처기업에 대한 정의를 재정리하였다. IT 벤처기업은 시장 환경의 변화가 매우 심하고, 제품에 대한 라이프사이클 또한 짧으므로, 기업의 성과관리가 매우 중요한 요인으로 작용할 것이다. 특히, 성과관리는 단기적인 차원이 아닌 중·장기적인 차원에서 전략적으로 지속되어야 하며, 이를 통한 지속적인 경쟁력 확보를 이룰 수 있도록 해야 한다(이현무 외, 2006). 또한, 같은 환경에 있는 IT 벤처기업들과의 재무상태에 대한 상대적인 평가를 통해 해당기업에 보다 더 영향을 줄 수 있는 요인이 무엇인지를 평가해보는 일도 매우 중요한 의미가 될 수 있을 것이다. 많은 연구자들이 벤처기업들의 특성을 이용하여 경영성과에 영향을 주는 요인들에 대해 연구하고 있으며, 특히 DEA라는 비모수적 접근법을 통해 IT 벤처기업의 재무자료를 이용하여 상대적인 효율성을 평가하는 것은 의미 있는 부분이 될 것이다.

### DEA

#### DEA의 개념

DEA(Data Envelopment Analysis)는 선형계획법의 응용 기법으로, 다수의 산출물을 생산하기 위해서 다양한 투입요소를 사용하고 있는 의사결정단위 (Decision Making Unit; DMU)들에 대한 어떤 DMU가 효율적인지, 효율성의 크기는 어느 정도인지 상대적인 효율성을 측정하고 평가하는 대표적인 비모수적 접근법이다. 단일 투입물과 단일 산출물의 효율성을 측정하기 위한 개념(Farrell, 1957)으로 시작하여, Charnes 등(1978)에 의해 CCR이라는 기본모형이 개발되었고, 이후 CCR 모형의 대안으로 많은 확장 모형이 개발되었는데, Banker 등(1984)에 의해 개발된 BCC 모형이 대표적이라 할 수 있으며, 아래의 식 중 (1)은 CCR 모형, (2)는 BCC 모형을 나타낸다.

$$\begin{aligned} & \min_{\theta, \lambda} \theta \\ & \text{subject to } \theta x_0 - X\lambda \geq 0 \\ & Y\lambda \geq y_0 \\ & \lambda \geq 0 \end{aligned} \quad (1)$$

$$\begin{aligned}
& \min_{\theta_B, \lambda} \theta_B \\
& \text{subject to } \theta_B x_0 - X\lambda \geq 0 \\
& Y\lambda \geq y_0 \\
& e\lambda = 1 \\
& \lambda \geq 0
\end{aligned} \tag{2}$$

DEA의 가장 기본모형이라고 할 수 있는 CCR 모형은 각 DMU들의 규모수익성을 고려하지 않고 효율성을 평가하는 방법으로 규모 효율성과 기술효율성을 구분하지 못한다. 반면에 BCC 모형은 CCR의 개선모형으로 규모효율성과 순수기술효율성을 분리하여 측정할 수 있다. 규모효율성(scale efficiency)은 규모의 불변을 가정한 CCR 모형의 효율성 값을 규모의 가변을 가정한 BCC 모형의 효율성 값으로 나누어줌으로써 얻을 수 있다. DEA에 의해 상대적으로 비효율적으로 판명된 DMU는 참조집합(reference set)을 통한 벤치마킹으로 효율성을 개선할 수 있다.

IT 벤처기업 운영에 있어서, 여러 가지 투입요소들을 개선함으로써 효율성을 추구하는 것이 용이할 것이므로 본 연구에서는 여러 가지 DEA 모형 가운데 현재의 산출요소를 유지하면서 투입요소를 최소화하는데 중점을 두는 투입요소 중심의 CCR, BCC 모형을 사용하여 IT 벤처기업의 상대적인 효율성을 평가하고자 한다.

CCR 모형은 효율성을 평가하는데 있어 해당 DMU의 규모효율성에 관련한 측면을 고려할 수 없는 단점이 있다. 이러한 CCR 모형의 대안으로 다양한 확장모형이 개발되었으며 대표적인 것이 BCC 모형이다. BCC 모형은 Banker 등(1984)이 개발한 것으로 CCR 모형의 경우, 일반적으로 규모의 수익(return to scale)이 일정하다고 가정하고 DMU들에 대한 효율성을 평가한 반면, BCC 모형에서는 효율적인 프론티어가 주어진 DMU들의 볼록 집합으로 구성되므로 규모의 수익성이 변동한다고 보고 있으며, DMU들의 평가에 있어 규모를 고려함으로써 효율성을 조금 더 의미 있게 평가할 수 있는 장점이 있다.

### DEA의 응용분야

DEA는 기업평가 및 신용평가 등의 영역에서 상대적인 효율성 및 생산성의 비교연구에 많이 사용되고 있으며, DEA 관련 많은 연구에서 신경망분석과 같은 인공지능방법과 로짓분석, 판별 분석, 그리고 군집분석과 같은 통계적인 방법을 이용한 다양한 시도가 계속되고 있다.

DEA를 이용하여 기업의 효율성을 분석한 기존의 연구를 살펴보면 김재준(2004)은 코스닥 기업의 수익성 결정 요소 및 코스닥 기업의 특성에 따른 영향력 차이를 알아보기 위해 DEA를 사용하여 효율성을 측정하고 결과 코스닥 기업의 원가효율성 정보가 수익성에 영향을 미친다는 것을 알아내었다. 이영찬

(2006)은 성과가 좋지 않은 기업을 효율적 프론티어에 위치시키는 부실관행 DEA를 신용위험평가에 도입하여 증화기법과 함께 사용해 봄으로써 부실관행 DEA 모형은 증화기법과 함께 사용할 때 신용위험문제에 매우 강력한 도구로 사용될 수 있음을 확인하였다. 구정옥 등(2006)은 코스닥에 등록된 소프트웨어기업의 생산성을 분석하기 위해 DEA 모형을 사용하였다. 국외의 관련연구를 살펴보면 Carlos et al.(2005)은 인터넷 기업을 대상으로 DEA 모형을 사용하여 효율성 분석을 실시하였으며, 통계적인 기법과 함께 사용하여 다양한 분석을 시도하였다. 이외에도, DEA는 은행과 같은 금융기관의 지점 간 효율성을 분석하거나, 호텔과 같은 서비스 산업에서 경쟁업체 간 상대적 효율성 분석에 자주 사용되고 있다.

### Support Vector Machine(SVM)

SVM은 Vapnik(1995)에 의해 개발된 것으로 입력공간과 관련된 비선형문제를 고차원의 특징공간의 선형문제로 대응시켜 나타내기 때문에 수학적 분석이 용이하다(Hearst et al., 1998). 특히 SVM이 최근 각광을 받는 이유는 대부분의 학습 알고리즘은 경험적 위험최소화에 기반하는 모형인데 반해, SVM은 구조적 위험최소화에 충실하므로, 분류문제에 있어 일반화 능력이 높아 많은 분야에서 응용되고 있다(Tay and Cao, 2001).

SVM은 분류 문제에 있어 훈련 데이터들을 서로 다른 두 개의 클래스로 분류할 때 분류의 기준이 되는 분리 경계면(hyper plane)을 학습알고리즘을 통해 찾고 클래스를 구분하는 최적분리경계면(maximum margin hyperplane)을 구하기 위해 분리경계면과 가장 가까운 점(support vector)과의 거리를 최대화한다. SVM이 두 개의 클래스 구분이 불가능 한 경우에는 커널함수를 이용하여 입력 자료를 고차원의 특징공간으로 사상시킴으로써, 고차원 공간에서는 선형 분리가 가능한 분리경계면을 생성할 수 있다. 그러므로, SVM에서는 커널함수의 역할이 매우 크다는 것을 알 수 있다.

SVM의 알고리즘을 살펴보면 분류문제에 있어서 학습 데이터를 이용하여 함수  $f: R^n \rightarrow \{\pm 1\}$ 을 추정하도록 하며, 분리되는 두 개의 클래스 중 A는  $x_i \in R^n, y_i = 1$ 로, B는  $x_i \in B, y_i = -1$ 로 표시한다. 학습 데이터가 선형으로 분리 가능하면 다음과 같이 나타낼 수 있다.

$$w \cdot x_i + b \geq +1, \forall x_i \in A \tag{3}$$

$$w \cdot x_i + b \leq -1, \forall x_i \in B \tag{4}$$

여기서,  $x$ 는 입력벡터,  $w$ 와  $b$ 는 분리경계면을 결정하는 모수로,  $w$ 는 가중치 벡터이고,  $b$ 는 바이어스이다.

식 (3)과 (4)를 이용하여, 아래의 식(5)를 도출할 수 있다.

$$y_i(w \cdot x_i + b) \geq 1, \forall x_i \in A \cup B \quad (5)$$

최대 마진분류자는 최대분리경계면을 가지고 데이터를 최적화하는데, 결국 식(6)의 최적화문제가 된다.

$$\min_{w,b} \frac{w \cdot w}{2} \quad (6)$$

$$s.t. \quad y_i(w \cdot x_i + b) \geq 1$$

최종적으로 최적 의사결정 분리경계면은 식(7)과 같다.

$$f(x, \alpha_i, b) = \sum y_i \alpha_i (x \cdot x_i) + b \quad (7)$$

여기서  $\alpha_i$  와  $b$  는 분리경계면을 결정하는 모수이며,  $x$  는 학습용 데이터,  $x_i$  는 서포트벡터를 나타낸다.

학습 데이터가 비선형인 경우에는 입력변수를 고차원의 특징공간으로 이동시켜 비선형 문제를 선형문제로 근사시킬 수 있으며, 비선형 분류문제에서는 식(8)과 같이 나타낸다.

$$f(x, \alpha_i, b) = \sum y_i \alpha_i K(x, x_i) + b \quad (8)$$

식 (8)에서  $K(x, x_i)$  를 커널함수라 한다. 커널함수는 원래 데이터를 고차원공간으로 사상시킴으로써 특징공간 내에 선형으로 분리 가능한 입력 데이터 셋을 만들며, 일반적으로 사용되는 커널함수는 다음과 같다.

다항식 커널:  $K(x, x_i) = (x \cdot x_i + 1)^d$

가우시안 RBF:  $K(x, x_i) = \exp(-\frac{1}{\delta^2}(x - x_i)^2)$

## 연구모형

본 연구는 전반적으로 높은 기술력은 가졌으나, 급변하는 환경 및 경쟁적인 기업환경에 의해 다수가 불안정 상태에 있는 IT 벤처기업들의 특성을 효율성 측면에서 분석하고자 한다. 이에 대표적인 비모수분석법인 DEA 를 사용하여 연구 대상기업의 효율성을 평가하고, 로짓모형을 통해 효율기업을 특징짓는 재무변수의 특성을 파악하고자 한다. 그리고, 효율기업 예측을 위해 우수한 기계학습기법으로 주목 받고 있는 SVM 모형을 적용하고자 한다. 최종적으로, 로짓, SVM, 판별분석방법 등 각 기법들에 대한 효율기업 예측 적중률을 비교해 보고자 한다.

연구 대상기업들의 효율성을 평가하기 위해, 먼저, DEA 모형을 이용하여 IT 벤처기업들을 효율기업과 비효율기업으로 분류한다. DEA 모형을 통해 기업의 효율성 분석을 수행한 많은 연구들(구정옥 외, 2006; Carlos et al., 2005; 김재준, 2004)이 있지만, DEA 모형은 분석대상 기업들의 상대적 효율성을 분석하기 때문에, 특정기업의 효율성을 단독으로 평가하기 어려운 단점이 있다. 이러한 문제점을 보완하는 방안으로 DEA 모형을 통해서 획득한 기업의 효율성 정보

를 이용하여 로짓모형을 학습시킬 수 있다.

IT 벤처기업의 재무정보와 효율성정보를 이용하여 로짓모형을 개발하게 되면, IT 벤처기업의 효율성을 평가하는데 유의한 재무정보를 찾을 수 있을 뿐만 아니라 DEA 모형을 통해 알 수 없었던 효율기업을 특징짓는 재무변수의 특성들을 파악할 수 있다. 효율기업과 비효율기업을 대상으로 단변량분석인 독립대응표본 t-test 를 수행하게 되면, 효율적 기업과 비효율적 기업을 구분해 주는 유의한 재무정보를 파악할 수 있다. t-test 에 의해 선정된 유의한 변수를 대상으로 로짓모형의 stepwise 방법을 이용함으로써 최종적으로 IT 벤처기업의 효율성 평가를 수행할 수 있는 평가모형을 구축할 수 있다. 이 평가모형을 이용하면, 모형의 구축 후에도 개별기업의 효율성을 예측할 수 있으며, 각 기업의 효율성에 따른 재무정보의 설명기능을 추가적으로 얻을 수 있다. t-test 및 로짓모형을 이용하여 도출된 변수들로 SVM, 판별분석등에 적용하여 실험결과를 비교분석하고자 한다. 본 연구의 연구모형은 <그림 1>과 같다.

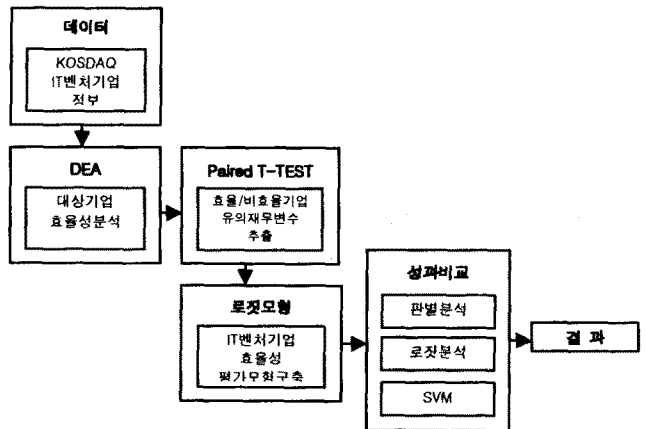


그림 1. 연구모형

## 실증분석

### 자료수집

본 연구에서는 코스닥에 등록된 기업 중 IT 벤처기업들의 상대적 효율성을 평가하기 위해 코스닥에 등록된 IT 벤처기업 93 개 업체를 선정하였으며, 그 중 연구 목적에 부합하지 않는 22 개 업체를 제외한 71 개 업체를 대상으로 분석을 실시하였다. 대상기업의 기업정보는 한국신용평가정보에서 제공하는 2005 년 기업공시자료를 활용하였다(<http://www.kisvalue.com>).

### DEA 효율성 평가

DEA 분석에 있어서는 투입요소와 산출요소를 선정하는 것이 매우 큰 이슈이다. 만약, 일부 기업에게 유리한 요소들을 투입요소 및 산출요소로 선정했을 경우, 다른 비교 대상보다 상대적으로 더 효율적이

라는 결과를 도출해낼 수 있기 때문이다. 본 연구에서는 DEA 를 이용한 소프트웨어 기업연구에서 사용된 투입요소 및 산출요소와, 여러 제조업에서 사용된 투입물 및 산출물을 토대로 IT 벤처기업의 특성을 잘 반영할 수 있는 투입요소와 산출요소를 선택하였다. 구정옥 등(2006)은 총자산, 직원수, 매출원가, 판매비와 관리비를 투입요소로, 매출액을 산출요소로 채택하여 국내 KOSDAQ 에 상장된 소프트웨어 기업에 대한 효율성을 평가하였다. 김강정(2005)도 국내 소프트웨어 산업의 효율성 평가에 투입요소로 총자산, 종업원 수, 매출원가 등을 사용하고, 매출액을 산출요소로 사용하였다. 아래의 <표 1>은 기존연구에서 사용된 투입요소 및 산출요소이다.

표 1 - DEA 에 사용된 투입요소 및 산출요소

연구자	투입요소	산출요소	적용분야
이경재 외(2007)	자본금, 유형자산, 급여, 광고선전비	순방문자수, 매출액	인터넷 기업
구정옥 외(2006)	총자산, 직원수, 매출원가, 판매비와 관리비	매출액	KOSDAQ S/W 기업
김건식(2006)	운영자산, 종업원수, 투입비용	외부매출, 내부매출	국내 SI 기업
김강정(2005)	총자산, 종업원수, 매출원가	매출액	국내 S/W 기업
김재준(2004)	노동, 유형고정자산, 연구개발비	매출액	KOSDAQ 제조기업
송동섭(2003)	노동, 자본, 중간재	매출액	국내 섬유산업
송동섭 외(2003)	노동, 자본, 중간재	매출액, 주가, 경상이익	국내 제조업
홍봉영(2003)	종업원수, 총자산, 총비용	매출액	국내 섬유산업
Barua et al.(2004)	IT 자본, NIT 자본, 종업원수, 사업연수	매출액, 순이익	닷컴기업

연구대상이 IT 벤처기업임을 고려하여, 본 연구에서는 기존 연구를 기반으로 종업원 수, 총자산, 개발비, 판매비와 관리비 등 총 4 가지의 투입요소와 산출요소인 매출액으로 연구를 진행하였다. 아래의 <표 2>은 본 연구에서 사용된 DEA 투입요소와 산출요소의 기초통계량이다.

표 2 - DEA 에 사용된 투입요소와 산출요소의 기초통계량 (단위: 명, 천원)

		평균	표준편차	최대값	최소값
투입물	종업원수	106.44	10.452	679	21
	총자산	37,625,408	3,841,019	203,220,181	7,598,282
	개발비	1,840,776	377,894.4	21,112,331	3,000

	판매비와 관리비	6,452,054	664,413.0	35,840,244	742,076
산출물	매출액	18,645,911	1,406,512	53,557,671	3,600,095

71 개의 IT 벤처기업 DMU(Decision Making Unit)에 대해 CCR 모형과 BCC 모형을 적용하여 효율성을 분석한 결과, CCR 모형에서 효율성이 1 인 DMU 는 모두 7 개였으며, 이 7 개의 기업은 나머지 64 개의 기업에 비해 상대적으로 효율적인 기업임을 알 수 있다. DMU 7, 12, 33 과 같은 효율성 점수가 0.2 미만인 DMU 들은 종업원 수 또는 총자산과 같은 투입요소의 수치는 높는데 반해 매출액이 상대적으로 낮다는 특징이 있다. 즉, IT 벤처기업의 경영문제에서 인적 자원과 금전적인 부분에서의 투자가 실제 기업수익으로 환원되지 않고 있다는 것을 알 수 있다. BCC 모형에서 효율성이 1 인 DMU 는 모두 16 개로 CCR 모형에서 9 개가 추가되었다. CCR 효율성이 1 이 아닌 DMU 들의 경우 비효율에 대한 문제가 기업 내부 경영상의 문제인지 또는 기업 규모에 있어 상대적으로 불리한 것인지 BCC 효율성을 통해 재검토 해야 한다. CCR 모형에서 비효율 기업으로 평가되었던 일부 DMU 들이 BCC 모형에서는 효율적인 DMU 로 선정되었다. 이 중 DMU 18 과 58 은 BCC 효율성이 1 이지만, CCR 효율성은 각각 0.3806 과 0.3321 로 매우 낮으며 이들의 규모효율성도 매우 낮다고 판단할 수 있다. 이 두 IT 벤처기업은 순수기술효율성은 우수하다고 할 수 있지만, 적은 인력으로 운영되는 등 규모에 있어 불리한 위치에 있다고 분석된다. CCR 효율성과 BCC 효율성이 모두 1 이 되지 않는 경우는, 규모에 대한 비효율과 운영에 대한 비효율을 상대적으로 평가될 수 있는데, DMU 17 과 33 의 경우는 CCR 효율성이 매우 낮음에도 불구하고 규모효율성은 0.9 보다 크므로 기업내부 운영은 잘 되고 있지 않으나 기업의 규모에 있어서는 유리한 위치에 있다고 볼 수 있다.

효율적인 DMU 의 경우, 자신은 스스로의 효율성 참조 집합이 되지만, 비효율 DMU 의 경우는 비효율을 개선하기 위해서 자신을 제외한 다른 DMU 들을 참조하게 된다. 비효율 DMU 들에 대한 참조집합의 참조빈도를 보면, DMU 31 이 가장 많이 참조되었음을 알 수 있다. DMU 31 은, CCR 모형에서 59 회, BCC 모형에서 36 회가 참조되었으며, 관련업계에서는 이 기업이 해당 투입변수 및 산출변수를 기준으로 평가할 때 가장 효율적인 기업이라고 판단할 수 있다. 추가적으로 투입변수의 조합을 상이하하게 하여 분석을 실시해 본 결과, 각각의 단일 투입요소에 따른 효율성 보다 투입요소를 둘 이상 조합하여 구한 효율성이 더 좋게 나타났다. 기업의 효율성에 대한 연구에서, 재무정보의 여러 가지 조합을 통해 효율성에 대한 추이를 볼 수 있을 것이며, 이러한 영향력을 통해 해당 기업의 효율성을 제시해줄 수 있는 재무정보의 구분이 가능할 것이다.

### 로짓모형을 이용한 DEA 그룹의 특성분석

DEA 분석을 통해 연구 대상 기업들 중 상대적으로 효율적인 기업들을 선별할 수 있었다. 그러나 DEA의 경우, 대상 기업들의 상대적인 효율성 판별에는 유용하나, 특정 기업의 효율성을 단독으로 판단하는 문제에 있어서는 무리가 있으므로, DEA를 통해 획득한 기업의 효율성 정보를 이용하여 로짓모형에 적용해 보고자 한다. 이때 사용되는 기업의 재무변수는, 효율적 기업과 비효율적 기업을 대상으로 하여 독립대응표본 t-t-test를 먼저 실시하여, 통계적으로 유의하다는 결론을 얻은 것들이다. 본 연구에서는, IT 벤처기업의 영향요인에 대한 많은 선행연구들 중, 특히 재무정보를 이용한 선행연구를 참조하였다. 총 92개의 재무변수를 토대로 하여, 효율기업과 비효율기업에 대한 독립대응표본 t-test를 통해 유의한 변수들을 선별하여 로짓분석을 실시하였다. 로짓모형의 stepwise method를 이용하여 변수 2개가 유의한 변수로 판명되었으며, 이 변수들을 이용하여 판별분석 및 SVM에도 이용하였다. 로짓모형을 이용한 결과는 아래의 <표 3>과 같다.

표 3 - 로짓모형 선택변수

변수명	B <sup>1)</sup>	S.E <sup>2)</sup>	Wald <sup>3)</sup>	유의확률
총자본회전율	5.818	1.696	11.771	0.001***
종업원 1인당 매출액	0.000	0.000	4.581	0.032**
상수	-5.990	1.475	16.490	0.000

1)로짓모형의 계수 2)표준오차 3)Wald 통계량  
 \*\*: 5% 수준에서 유의함. \*\*\*: 1% 수준에서 유의함

<표 3>은 로짓분석의 회귀식에 포함된 재무변수 정보를 담고 있으며, 총자본회전율은 유의수준 1%에서, 종업원 1인당 매출액은 유의수준 5%에서 유의하다는 결론을 얻었다. 이 두 재무정보는, 기업의 자산운영, 그리고 기업의 생산성에 관련된 재무정보인 총자본회전율과 종업원 1인당 매출액이 효율기업이라는 판단을 가능하게 하는 영향력 있는 재무변수임을 확인할 수 있다.

### SVM 모형구축

본 연구에서는, SVM에서 비선형함수로 많이 사용되고 있는 가우시안 RBF 함수를 사용하였으며, 모수인 C와  $\delta^2$ 에 따라 성과가 달라진다. 최적화 모형을 구축하기 위해서 모수인 C와  $\delta^2$ 의 조절이 필요하며, 본 연구에서 모수 C는 20, 40, 60, 80, 100의 5가지 경우로, 그리고,  $\delta^2$ 는 1, 25, 50, 75, 100의 경우로 나누어 실험하였다. 실험은 Libsvm-2.83으로 실행하였으며, 실험 결과 전체 데이터셋에 대해 같은 값으로 묶인 구간이 많이 발견되었다. 결과적으로

$\delta^2=25$ 일 때, 데이터셋 1, 2, 3이 비교적 높은 예측율을 보였으며, 이 때의 C 값은 데이터셋 1이 60, 데이터셋 2와 3이 20이었다. 또한,  $\delta^2=50$ 일 때는, 데이터셋 5가 비교적 높은 예측율을 보였으며, C 값은 60이었다. 데이터셋 4는,  $\delta^2=100$ , C 값이 20일 때 높은 예측율을 보였다.

### 효율기업과 비효율기업의 예측성과

본 연구에서는 로짓모형의 성과 분석 및 판별분석, SVM에 대한 성과분석을 위해 데이터를 학습용과 검증용으로 분류하였다. 또한, 로짓분석, 판별분석, SVM 실험에 대해 데이터 전 구간에 걸친 검증을 위해서 모두 70개의 데이터에 대해 5-fold cross-validation을 실시하였으며, 아래의 <표 4>는 IT 벤처기업의 효율기업 비효율기업 예측에 사용된 세 기법간의 예측율 결과를 제시한 것이다.

표 4 - IT 벤처기업의 효율기업 및 비효율기업 예측율(%)

	판별		로짓		SVM	
	학습용	검증용	학습용	검증용	학습용	검증용
Set1	82.1	78.6	83.9	85.7	80.4	85.7
Set2	78.6	92.9	82.1	92.9	78.6	92.9
Set3	87.5	78.6	85.7	71.4	82.4	78.6
Set4	83.9	78.6	85.7	78.6	82.1	78.6
Set5	82.1	78.6	85.7	85.7	85.7	85.7
평균	82.8	81.5	84.6	82.9	81.8	84.3

IT 벤처기업에 대한 효율기업과 비효율기업 예측성과는 판별, 로짓, SVM 방법 간에 차이가 있음을 알 수 있다. 예측율의 성과는 SVM, 로짓분석, 판별분석의 순으로 높게 제시되었다.

### 결론 및 향후연구

본 연구의 목적은 코스닥에 등록된 소프트웨어 서비스 관련 IT 벤처기업을 대상으로 해당 기업들의 상대적인 효율성을 평가하고, 효율기업과 비효율기업을 특징짓는 주요 재무자료의 특성을 고찰하여 최종적으로 특정기업의 효율성을 단독으로 평가할 수 있는 모형을 구축하는데 있다.

코스닥에 상장된 소프트웨어 서비스 관련 IT 벤처기업을 대상으로 한국신용평가정보에서 제공하는 2005년 기업재무정보를 바탕으로 연구를 진행하였으며, 주요 연구방법은, 비모수분석기법인 DEA를 사용하였으며 이는 연구대상 기업들의 상대적 효율성을 평가하기 위함이다. DEA를 이용한 효율성 분석 결과를 바탕으로 이분화 된 효율기업과 비효율기업을 대상으로 독립대응표본 t-test를 실시하였다. t-test를 통

해 통계적으로 유의한 재무정보를 파악한 후, 로짓 모형의 stepwise 방법으로 최종적인 IT 벤처기업 효율성 평가 모형을 구축하였다. DEA 가 여러 기업들에 대한 상대적인 효율성 평가에는 유용하지만, 특정기업의 효율성에 대한 단독 평가에는 다소 어려움이 따른다. 본 연구에서의 DEA-logit 통합모형을 통해, 특정기업에 대한 평가가 가능하다. 분석 결과, 총자본회전율이나 종업원 1 인당 매출액 등 기업자본의 효율적 운영이나, 종업원의 생산성이 IT 벤처기업의 효율성을 평가하는 주요 재무변수가 됨을 알 수 있었다. 효율기업의 분류문제에 대해, SVM, 로짓분석, 판별분석의 예측율을 비교했을 때, 전반적으로 SVM의 예측율이 다른 두 기법에 비해 우수함을 알 수 있었다.

본 연구에서는 연구대상 표본수가 한계점으로, 다양한 분석적용이 어려웠다. 향후 연구에서는, IT 벤처기업에 적용하여 구축한 효율성 평가 모형을 다른 산업분야 또는 다른 나라의 IT 벤처기업에 적용하여 일반화할 수 있도록 개선되어야 할 것이다.

## 참고문헌

- [1] 구정옥, 김강정, 박정원(2006), "KOSDAQ 소프트웨어기업의 생산성분석," *재무와 회계정보저널*, 제 6 권 제 1 호, pp.27-51.
- [2] 김강정(2005), "국내 소프트웨어 산업의 효율성 분석," *생산성논집*, 제 19 권 제 1 호, pp.95-111.
- [3] 김건식(2006), "자료포괄분석과 생산성지수분석을 이용한 국내 SI 기업의 효율성 분석," *한국SI 학회지*, 제 5 권 제 1 호, pp.1-16.
- [4] 김우식, 이광조(2003), "KOSDAQ 소프트웨어 기업의 회계정보를 이용한 경영 효율성 평가에 관한 연구," *회계정보연구*, 제 21 권, pp.357-378.
- [5] 김재준(2004), "코스닥기업의 DEA 효율성과 수익성의 관련성 연구," *세무회계연구*, 제 14 권, pp.153-170.
- [6] 서종상, 최현돌(2006), "코스닥시장에서 벤처기업과 일반기업의 재무적 및 주식 가치적 특성", *Journal of Business Research*, Vol.21 No.1, pp.317-345.
- [7] 송동섭(2003), "중소호텔산업의 효율성 제고와 성과측정지표의 적용가능성," *중소기업연구*, 제 25 권 제 1 호, pp.199-223.
- [8] 송동섭, 김재준(2000), "DEA 모형을 이용한 제조업의 효율성 분석에 관한 연구," *회계정보연구*, 제 14 권, pp.127-152.
- [9] 이경재, 김재전, 조건(2007), "DEA 모형을 활용한 인터넷 기업의 효율성 평가," *대한경영학회지*, 제 20 권 제 1 호, pp.109-136.
- [10] 이영찬(2006), "부실관행 DEA 를 이용한 신용위험 평가의 정확도 향상," *대한경영학회지*, 제 19 권 제 5 호, pp.183-203.
- [11] 이장우(1998), "성공 벤처기업 특성에 관한 탐색적 연구," *벤처경영연구*, 제 1 권 제 2 호, pp.69-95.
- [12] 이진주(1998), "벤처경영의 특성과 발전과제," *벤처경영연구*, 제 1 권 제 1 호, pp.213-244.
- [13] 이춘우(2005), "벤처기업의 개념적 특성과 본질에 관한 연구," *경영학연구*, 제 34 권 제 2 호, pp.315-348.
- [14] 이현무, 강민철(2006), "IT 벤처기업의 경영성과 영향요인에 관한 연구," *대한경영학회지*, 제 19 권 제 3 호, pp.1059-1092.
- [15] 조준희(2003), "벤처기업의 우량 및 비우량 재무 특성-코스닥 등록기업을 중심으로," *전산회계연구*, 제 2 권 제 1 호, pp.221-234.
- [16] 홍봉영(2003), "DEA 를 이용한 국내 섬유산업의 효율성 및 생산성변화 분석," *산업조직연구*, 제 11 권 제 38 호, pp.61-96.
- [17] Banker, R. D., A. Charnes and W. W. Cooper (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science*, Vol.30, pp.1078-1092.
- [18] Barua, A.P.L.Brockette, W.W. Cooper, H.Deng, B.R. Parket, T.W. Ruefli and A. Winston (2004), "DEA evaluations of long-and short-run efficiencies of digital vs. physical product "dot com" companies," *Socio-Economic Planning Sciences*, Vol.38 No.4, pp.233-253.
- [19] Carlos, S.C., Yolanda, F.C., and Cecilio, M.M. (2005), "Measuring DEA efficiency in Internet companies," *Decision Support Systems*, Vol.38, pp.557-573.
- [20] Chang, C.-C. and Lin, C-J.(2007), LISVM: a library for support vector machines, Technical Report, Department of Computer Science and Information Engineering, National Taiwan University, Available at <http://www.csie.edu.tw/~chlin/papers/libsvm.pdf>.
- [21] Charnes, W. W. Cooper and E. Rhodes(1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operation Research*, Vol.2, pp.429-444.
- [22] Farrell, M. J. (1957), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society*, Vol.120, pp.253-281.
- [23] Hearst, M.A., Dumais, S.T., Osman, E., Platt, j. and Scholkopf, B.(1998), "Support vector machines," *IEEE Intelligent System*, Vol.13, No.4, pp.18-28.
- [24] Tay, F.E.H., and Cao, L.J.(2002), "Modified Support Vector Machines in Financial Time Series Forecasting," *Neurocomputing*, Vol.48, pp.847-861.
- [25] Vapnik, V.(1995), *The Nature of Statistical Learning Theory*, Springer.