

기업 인적자원 관련 변수를 이용한 기업 신용점수 모형 구축에 관한 연구*

이영섭¹⁾ 박주완²⁾

요 약

본 논문의 목적은 기업 신용점수에 영향을 미치는 기업 인적자원 요소들을 찾아서 기업 신용점수 모형을 구축하는 것이다. 모형 구축을 위해 사용된 자료는 2005년 한국직업능력개발원의 인적자본 기업패널 (Human Capital Corporate Panel, HCCP) 설문조사 자료와 한국신용평가(주)의 KIS-신용평점모델에서 생성된 기업 신용점수이다.

모형 구축을 위한 독립변수는 McLagan (1989)의 '인적자원 바퀴모델'을 토대로 인적자본 기업패널 설문조사 문항을 선택하여 사용하였으며, 종속변수로는 기업 신용평가 점수를 사용하였다. 또한 기업 인적자원 관련 변수를 이용한 기업 신용점수 모형 구축을 위해 로지스틱 회귀모형을 사용하였다.

모형 구축 결과 최종적으로 선택된 변수는 22개였다. 영역별로 세분화해서 살펴보면 대분류 기준으로 HRD 영역은 6개, HRM 영역은 15개, 기타 1개이고, 중분류 기준으로 개인개발 2개, 경력개발 2개, 조직개발 2개, 조직직무설계 1개, 인적자원계획 4개, 정보체계 2개, 보상 및 장려 6개, 복지후생 1개, 노사관계 1개, 기업규모 1개가 선택되었다.

구축된 모형을 평가하기 위하여 10등급 교차타당성 분석을 통한 오분류율, G-mean은 각각 30.81, 68.27이었다. 그리고 반응율은 가장 좋은 십분위가 가장 나쁜 십분위보다 약 6.08배가 크고 점차 감소하는 경향을 보이고 있다. 그러므로 구축된 모형은 기업 인적자원 관련 변수를 이용해 기업 신용점수를 측정하는데 적당한 모형이라는 결론을 내릴 수 있다.

주요용어: 데이터마이닝, 로지스틱 회귀모형, 신용점수, 인적자원관리, 인적자원개발.

1. 서론

현재 기업은 우수한 인적자원을 어떻게 채용하고 양성하며 효율적으로 관리할 것인가의 문제에 당면하고 있다. 특히 경제 환경이 세계화되고 경쟁이 더욱 치열해짐에 따라, 우수한 인적자원 확보는 기업의 시급한 과제가 되고 있으며 (송창용 등, 2006a), 인적자원은 회사의 무형적 자산 또는 지식자산의 하부구조적 자산으로서, 기업의 장기적 수익 창출 및

* 본 연구는 한국과학재단 목적기초연구 (과제번호: R01-2004-000-10689-0) 지원으로 수행되었음.

1) (100-715) 서울특별시 중구 필동 3가 26번지, 동국대학교 통계학과, 부교수

E-mail: yung@dongguk.edu

2) (100-715) 서울특별시 중구 필동 3가 26번지, 동국대학교 통계학과, 박사과정

E-mail: jwan0217@naver.com

성과 제고의 핵심적 가치와 위치를 차지하고 있다. 이에 따라 미국·유럽의 주요 기업들은 인적자원의 가치를 계량화하려는 노력을 30년 이상 지속 중이다(강혜영, 2003).

인적자산에 대한 개별 기업의 노력과 성과로 최근 들어 국가 차원의 회계 기준화를 위한 공동 노력이 활발하다. 또한 이는 인적자산을 확대하기 위한 활동으로 볼 수 있는 인적자원 활동의 효과 측정으로 이어지고 있다. 현재, 인적자원의 양과 질, 인적자원개발 및 관리 관행에 대한 다양한 측정(Huselid 등, 2005; Becker 등, 2001)이 시도되고 있다. 그러나 인적자원 활동에 대한 측정도구 및 모형의 개발 노력은 아직 시작 단계에 불과하다. 그러므로, 기업 내부의 인적자원 축적 과정과 기업 가치를 포함한 경영 성과의 인과관계를 밝히는 기초 작업을 통해 인적자본지수(Human Capital Index, HCI)를 개발하는 것이 국가적인 인력개발 정책의 수립이나 기업의 인적자원관리 전략 수립에 중요한 과제로 부각되고 있다(김용민, 2006). 해외 선진기업은 인적자원 활동의 투자수익률(Return On Investment, ROI)을 측정하여 인적자원 활동이 경영 성과의 향상과 직결된다는 입증을 해오고 있는 지 오래이며, 이러한 활동이 뒷받침되어 많은 기업들이 적극적으로 인적자원 활동에 나서고 있다(이성, 2003). 그리고 기업 신용평가 점수와 인적자원 활동 사이에 어느 정도 상관성 있다는 것도 여러 연구를 통해 제시되고 있다. 송창용 등(2006b)은 중소기업의 인적자원 관련 활동과 중소기업의 신용수준의 상관성 분석을 통해 100인~299인 규모의 기업에서 인적자원 관련 활동이 기업 신용도와 유의미한 상관관계가 있음을 보였다.

그동안 다양한 종류의 기업 관련 변수들을 이용하여 기업 신용평가 모형을 구축하는 여러 가지 연구들이 있었지만 기업 인적자원 요소를 활용하여 기업 신용평가를 한 적은 없었다. 그러므로 본 논문에서는 기업 내부의 인적자원 요소와 기업 가치를 포괄하는 경영 성과의 결과 간의 인과관계를 밝히는 기초 작업의 일환으로, 기업의 신용도는 기업의 경쟁력에 영향을 주고 있고 그러한 기업의 경쟁력은 기업의 인적자원과 상관이 있다는 점에 착안하여, 데이터마이닝(data mining) 기법 중 로지스틱 회귀모형(logistic regression model)을 이용하여 기업 인적자원 관련 변수를 이용한 기업 신용점수 모형을 구축하고자 한다. 로지스틱 회귀모형을 사용한 이유는 정확성이 우수하고, 구축 과정이 용이하고 해석하기가 쉬우며, 과대 적합(over-fitting)할 가능성이 적고, 오차를 최소화하는 선형적인 관계를 찾는 데 매우 우수한 기법이기 때문이다(이영섭, 2003).

기업 인적자원 관련 변수를 이용한 기업 신용점수 모형 구축을 위해 2005년 한국직업능력개발원의 인적자본 기업패널(Human Capital Corporate Panel, HCCP) 1차년도 조사 자료와 한국신용평가정보(주)의 KIS-신용평점모델에서 생성된 2004년 기업 신용평가 점수를 이용하였다. 여기에서 인적자본 기업패널 자료는 2005년에 조사되었지만 조사된 내용들은 2004년도의 기업 인적자원에 관한 사항들이기 때문에 2004년 기업 신용평가 점수와 자료를 합쳐서 사용해도 자료 통합 시점 상에는 큰 문제가 없는 것으로 사료된다. 그리고 구축된 모형의 비교 및 평가는 오분류율, G-mean, 반응율(percent of response)을 이용하였다.

논문의 구성은 다음과 같다. 2절에서는 예측 모형 구축을 위한 알고리즘과 모형 평가 방법을 설명하였으며, 3절은 기업 신용점수 모형을 구축하기 위한 연구 대상, 변수 선정, 모형 구축 과정에 대해 설명하였고, 4절은 실제 자료를 이용하여 모형을 구축하고 평가한 결과이다.

2. 모형 구축을 위한 알고리즘 및 모형 평가

본 논문에서 사용할 데이터마이닝 예측 모형 구축을 위한 알고리즘은 종속변수의 범주가 0과 1 두개일 때 범주 1이 될 확률을 구하기 위하여 예측을 수행하는 로지스틱 회귀모형이다. 분석을 위해 사용된 로지스틱 회귀모형에 대해 간략히 소개하면 다음과 같다.

로지스틱 회귀모형은 종속변수 Y_i 가 이진형(binary type) 또는 순서형(ordinal type)인 경우 반응함수 $E(Y_i|x_i's)$ 는 $x_i's$ 가 증가함에 따라 값이 1로 서서히 수렴하는 모형으로, 종속변수가 1이 될 확률인 $\Pr(Y = 1|x_i's)$ 를 예측하는 모형이다. 즉, 로지스틱 회귀모형은 관심을 갖고 있는 쪽의 결과로 나타나는 반응확률과 독립변수 사이의 관계를 연구하기 위하여 사용되는 회귀모형이다.

모형 구축 시 로지스틱 회귀모형이 많이 선호되고 있다. 그 이유는 첫째 모형 구축이 올바르다면 로지스틱 회귀모형은 정확성이 우수하고, 둘째 구축 과정이 용이하고 해석하기가 쉬우며, 셋째 과대 적합(over-fitting)할 가능성이 적고, 오차를 최소화하는 선형적인 관계를 찾는 데 매우 우수한 기법이기 때문이다(이영섭, 2003). 로지스틱 회귀모형에 대한 자세한 이론적 설명은 Hosmer와 Lemeshow (2000)를 참고하기 바란다.

최적의 모형을 얻기 위해서는 여러 모형을 비교, 평가하여 하나의 모형을 선택하고 선택된 모형이 다른 모형에 비해 우수하다는 것을 입증해야 한다. 이 때 고려되어야 할 사항 중 가장 우선적인 것은 구축된 모형이 얼마나 예측과 분류에서 뛰어난 성능을 가지는지 알아보는 것이다. 결과적으로 모형 평가란 예측을 위해 만든 모형이 다른 임의의 모형보다 우수한지, 서로 다른 모형들 중에서 어느 것이 가장 좋은 예측력을 보유하고 있는지를 비교, 분석하는 과정이다.

일반적인 모형 평가의 기준은 모형의 설명력의 척도인 결정계수(R-square), AIC(Akaike Information Criterion), BIC(Bayesian Information Criterion) 등의 통계량을 통해 측정될 수 있으며, 특히 종속변수가 범주형인 경우 오분류표를 통한 여러 가지 방법을 사용할 수 있다. 그리고 SAS E-miner에서는 모형 평가 시 일반적으로 사용되는 방법으로 모형의 성능을 오분류표에 의한 민감도(sensitivity)와 특이도(specificity)에 의해 판단하는 ROC도표, 어떠한 범주가 각 집단에서 평균적으로 가지는 빈도와 해당 집단 내의 범주에 대한 비율을 이용한 리프트도표, 반응율도표 등이 있다(강현철 등, 1999).

오분류율은 범주가 0과 1인 두 가지의 결과를 가질 경우 실제 0인 범주를 1 또는 실제 1인 범주를 0으로 오분류하는 확률을 말하는 것이다. 즉, 확률변수 X 를 0 또는 1의 값을 가지는 실제값이라고 하고 확률변수 E 는 0 또는 1의 값을 가지는 예측값이라고 할 때 오분류율은 다음과 같이 나타낼 수 있다.

$$\text{오분류율} = \Pr(X = 1, E = 0) + \Pr(X = 0, E = 1). \quad (2.1)$$

오분류율은 전체 자료를 얼마나 잘못 분류하는가의 문제이므로 그 값이 적을수록 좋은 모형이라는 결론을 내릴 수 있다.

반응율은 모형을 통해 사후확률을 구한 후 사후확률의 순서에 따라 전체 자료 세트를 정렬하고 이를 N 개로 등분한 후 각 집단에서 반응변수의 특정 범주에 대한 빈도를 계산하

여 다음과 같이 반응율을 계산한다.

$$\text{반응율 값} = \frac{\text{일정 } N\text{등분 내 범주 1빈도}}{\text{일정 } N\text{등분 내 전체 빈도}} \quad (2.2)$$

이렇게 구해진 반응율을 그림으로 나타낸 도표가 반응율 도표인데 점수가 높은 10 분 위수에서 높게 나타나다가 급격하게 감소하는 형태이거나, 그래프가 기준선으로부터 멀리 떨어져 있는 형태인 경우 비교적 좋은 예측 모형이다.

G-mean은 결과 범주가 0인 집단과 1인 집단을 동등하게 고려하는 측도로써 다음과 같이 측정될 수 있다.

$$G - \text{mean} = \sqrt{\frac{\text{실제 0, 예측 0 정분류 빈도}}{\text{실제 0 빈도}} \times \frac{\text{실제 1, 예측 1 정분류 빈도}}{\text{실제 1 빈도}}} \quad (2.3)$$

G-mean은 실제 범주 0인 집단에 대한 정확도와 범주 1인 집단에 대한 정확도의 기하평균으로 이해될 수 있다. 그러므로 G-mean의 값이 클수록 좋은 예측 모형이다.

최종적으로 구축된 모형은 10등급 교차타당성(10-fold cross validation)을 이용하여 오분류율, 반응율, G-Mean의 값으로 모형을 비교 및 평가하고자 한다.

3. 모형 구축

3.1. 연구 대상

본 연구의 대상은 2005년도 한국직업능력개발원 인적자본 기업패널 설문에 응답한 454개 기업체 중 종업원 규모 100인 이상이고 2004년 한국신용평가(주)의 기업 신용점수 자료가 있는 기업체를 모형 구축을 위한 대상으로 한정하였다. 454개의 설문응답 기업체 중 종업원 규모가 100인 이상이면서 신용점수가 있는 업체는 모두 412개였다. 모형 구축을 위한 대상을 100인 이상으로 한정하는 이유는 기업이 자체적으로 인적자원개발 활동을 하기 위해서는 어느 정도 회사의 규모가 있어야 하기 때문이다.

3.2. 변수 선정

기업 인적자원 수준 측정을 위한 기업 신용점수 모형 구축에 사용할 독립변수는 인적자원 활동 관련 항목들을 사용하고자한다. 이 때 독립변수인 인적자원 관련 활동 항목들은 인적자원바퀴(HR Wheel) 모델을 토대로 설정하였다. 그림 3.1의 인적자원바퀴모델은 McLagan (1989)이 제시한 것으로 기업 내 인적자원개발 실무자들에게 가장 일반적으로 수용되고 있는 것이다(김미숙 등, 2005).

그림 3.1에서 알 수 있듯이 인적자원 바퀴모델 구성 요소는 인적자원개발(Human Resource Development, HRD), 인적자원관리(Human Resource Management, HRM)로 구성되어 있다. 인적자원개발은 개인개발, 경력개발, 조직개발의 3개의 영역으로 나뉘어져 있으며 인적자원관리는 직무설계, 인적자원계획, 선발 및 임명, 인적자원정보체계, 보상 및 장려, 근로자 복지후생, 노조근로자 관계로 7개의 영역으로 구성되어 있다.

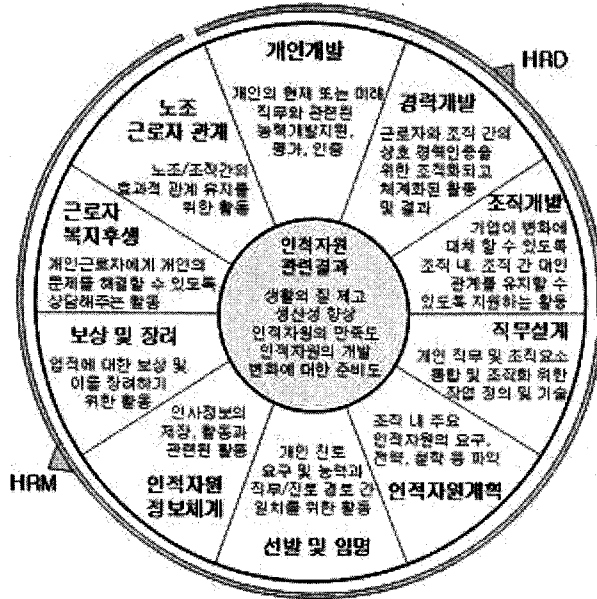


그림 3.1: 인적자원 바퀴(HR Wheel)모델

McLagan (1989)이 제시한 인적자원관리 영역과 인적자원개발 영역의 구체적인 활동 내용을 2005년도 한국직업능력개발원에서 실시한 인적자본 기업패널의 설문 항목들과 연계시켰다. 분석에 사용된 설문 문항은 총 77개로, 대분류 영역별로는 HRD 영역 29개, HRM 영역 47개, 기타 1개 문항이고, 중분류 영역별로는 개인개발 9개, 경력개발 12개, 조직개발 8개, 조직직무설계 7개, 인적자원계획 8개, 선발 및 임명 4개, 인적자원정보체계 11개, 보상 및 장려 11개, 근로자 복지후생 5개, 노사관계 1개, 기업규모 1개 문항이다.

표 3.1은 모형 구축에 이용하기 위해 McLagan의 인적자원 바퀴모델과 인적자본 기업패널 설문 항목을 연계한 문항들이다.

중속변수인 기업 신용점수는 2004년 한국신용평가(주)에서 조사한 자료를 이용하였다. 한국신용평가(주)에서 구한 기업 신용점수는 12개의 재무지표를 이용하여 산출한다. 이 때 이용하는 12개의 재무지표는 총자산, 매출액, 자기자본비율, 부채대비매출비율, 순영업자본회전율, 보조지표(유동비율, 영업활동 CF), 총자산순이익율, 금융비용부담율, 총자산증가율, 매출증가율, 총자산회전율, 매출채권회전율이다. 각 재무지표별로 점수가 산출되는 방식은 대상 기업군에서 지표별로 순위를 매기고 전체 분포에서 차지하는 상대적인 위치에 따라 점수가 주어지게 된다. 이렇게 모든 평가 요소들에 대한 점수가 계산되고, 이를 총합하면 최종 신용점수가 산출된다. 이러한 방식으로 최종적인 신용점수 값이 계산되면 평점이 속하는 구간별로 우수, 양호, 보통, 열위, 불량 등의 5단계로 평가된다. 구체적으로 신용점수가 80점 이상은 우수, 70점~79점은 양호, 55점~69점은 보통, 45점~54점은 열위, 44점 이하의 불량으로 나눌 수 있다. 신용점수 모형은 당해 연도 재무제표 기준에 근거하여 상

표 3.1: 모형 구축에 이용된 설문

대분류	중분류	소분류(설문문항)
인적 자원 관리 (HRD)	개인개발	집체식 사내 교육훈련 여부, 집체식 사외 교육훈련 여부, 인터넷 학습 여부, 우편통신 훈련 여부, 국내 연수 여부, 해외 연수 여부, OJT 프로그램, TFT 등 특별과제 팀 프로그램 여부, 개인 프로젝트 여부
	경력개발	국내 대학 등록금 지원 여부, 국내 대학원 등록금 지원 여부, 해외 대학원 학위 과정 지원 여부, 경력개발 제도 여부, 교육훈련 휴가제 여부, 학원 수강료 지원 여부, 사내공모제 여부, 멘토링 및 코칭 여부, 직무순환 여부, 자격수당 제도 여부(사내자격, 국가자격, 민간자격)
	조직개발	종업원 만족도 조사 여부, 승계계획(Succession Planning) 여부, HR 기능-성과 및 역량평가 여부, 균형평가표(BSC) 여부, 목표에 의한 관리(MBO) 여부, 다면평가 여부, 6-시그마 여부, QC(품질분임조) 여부
인적 자원 개발 (HRM)	조직직무 설계	직급 단순화 실시 여부, 직급 폐지 실시 여부, 직급 직책의 분리 여부, 직군별 차별화된 인사제도 여부, 전문직제 실시 여부, 직무기술서 실시 여부, 직무평가 실시 여부
	인적자원 계획	HR기능(인력계획 시스템 여부, 인력계획 컨설팅 여부), 채용의 전문성 여부(인성면접, 전문능력심사, 집단토론), HR업무담당전담조직유무, 정기적인 인력 계획수립 여부, HR 업무에 관한 관리자(임원, 팀장)의 전략 이해도 여부
	선발 및 임명	전체 공식 중 사내공모제를 통해서 충원 여부, 작업반 내 이동 여부, 작업반 간 이동 여부, 배치전환 기준 여부
	인적자원 정보체계	인사정보시스템 발전 단계, 인사정보시스템 여부(급여, 평가, 전환배치, 승진/승격, 교육훈련, 직무관리, 경력개발, 채용, 인사정보, 사내공모)
	보상 및 장려	호봉제 여부, 연봉제 여부, 직무급 여부, 직능급 여부, 개인 성과급 여부, 이익 배분 제도 여부, 팀 성과급 여부, 사업부 성과급 여부, 전사 성과급 여부, 종업원지주제 여부, 스톡옵션 여부
	복지후생	선택적 복리후생 여부, 복리후생 수준(사원 1년차, 과장 1년차, 부장 1년차), 인사정보시스템 가동 여부
	노사관계	노사관계
기타	기업규모	기업규모

대적으로 객관적·정량적으로 신용관리의 합리성을 확보하고 있다(한국신용평가, 2005).

본 논문에서는 로지스틱 회귀모형을 구축하기 위해 종속변수를 신용점수가 우수 및 양호 등급인 70점 이상을 “1”로하고 보통, 열위, 불량인 70점 미만을 “0”으로 변환하여 사용하였다.

3.3. 모형 구축 절차

본 논문의 목적은 신용평가 점수에 영향을 미치는 기업 인적자원 요소들을 찾아서 신용점수 모형을 구축하는 것이다. 모형 구축 절차는 그림 3.2와 같다.

첫째, 한국신용평가정보(주)의 2004년 기업 신용평가 점수 자료와 한국직업능력개발원의 2005년 인적자본 기업패널 설문 자료를 이용하여 모형 구축용 자료 세트를 구성하였다. 이 두 자료를 하나로 합쳐 종속변수로는 신용평가 점수를 사용하였고, 독립변수는 인적자본 기업패널 설문지에서 인적자원 활동 관련 설문 항목을 선정하여 이용하였다. 둘째, 모형 구축을 위한 변수를 선택하였다. 최초 77개의 변수(설문 문항)를 이용하여 로지스틱 회귀모형에서의 변수 선택 방법인 단계적 선택법(stepwise) 및 후진소거법(backward)으로 유의확률이 0.3 이상인 변수를 1차적으로 제거하였다. 그리고 최종 모형을 구축하기 위한 변수는 최적조합 선택법(score)을 이용해 선택하였다. 셋째, 최종 모형을 구축하고 평가하였다. 최적조합 선택법으로 선택된 변수를 이용하여 최종 모형을 구축하였으며, 이렇게 구축된 모형에 대한 평가는 10등급 교차타당성(10-fold cross validation) 방법을 이용하였다. 모형 구축에 사용된 알고리즘은 로지스틱 회귀모형이며 구축된 모형의 평가를 위해 오분류율, G-mean, 반응율을 이용하였다.

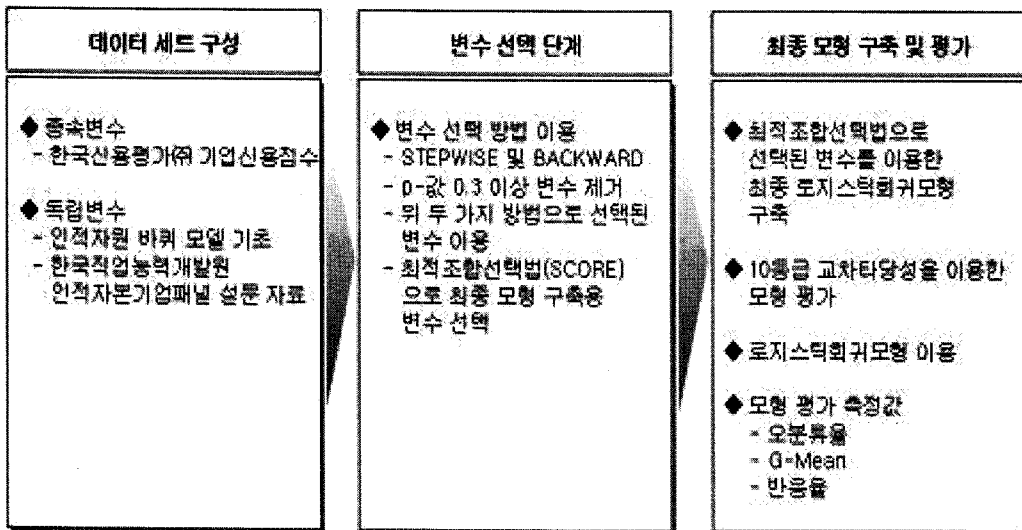


그림 3.2: 모형 구축 과정

4. 분석 결과

4.1. 변수 선택

기존 변수들을 결합하거나 조합함으로써 새로운 변수를 생성하는 여러 가지 방법이 있다. 이렇게 하기 위해서는 자료와 현업을 잘 이해하는 것이 중요하다. 자료를 추출하고 정제하는 일련의 과정이 모두 끝났다면 모형을 구축하기 위해 가장 영향력 있는 변수들을 선택해야 한다. Hosmer와 Lemshow (2000)에 의하면 변수 선택 시 각 변수에 대하여 단변량 로지스틱 회귀분석을 하도록 제안했지만 변수의 수가 아주 많을 경우 시간이 많이 걸리며 비효율적이다. 따라서 본 연구에서의 변수 선택 방법은 이영섭 (2003)에 제시된 방법을 이용하였다.

그림 3.2에서의 변수 선택 단계 중 첫 번째 단계를 이용하여 선택된 변수들은 표 4.1과 같다. 최초 77개의 모든 변수 중 로지스틱 회귀분석을 이용하여, 유의확률 0.3을 기준으로 단계적선택법 및 후진소거법을 이용해 변수를 선택한 것이다. 선택된 변수는 총 28개였다. 선택된 변수를 분류 영역별로 살펴본 결과 대분류 기준으로 HRD 영역 12개, HRM 영역 15개, 기타 1개이고, 중분류 기준으로 개인개발 5개, 경력개발 3개, 조직개발 4개, 조직직무 설계 1개, 인적자원계획 4개, 정보체계 2개, 보상 및 장려 6개, 복리후생 1개, 노사관계 1개, 기업규모 1개이다. 여기에서 유의확률이 아주 큰 0.3을 기준으로 한 이유는 1차 변수 선택에 있어서 아주 유의하지 않다고 판단되는 변수들만 우선 대략적으로 제거하기 위해서이다.

표 4.1: 1차 선택 변수-단계적 선택법 및 후진소거법

대분류	중분류	소분류(설문문항)
HRD	개인개발	집체식사내교육훈련 여부, 인터넷학습 여부, 우편통신훈련 여부, 국내연수 여부, 해외연수 여부
	경력개발	해외대학원학위과정지원 여부, 사내공모제 여부, 국가자격수당제도 여부
	조직개발	HR기능-성과및역량평가 여부, 균형평가표(BSC)여부, 6-시그마 여부, QC(품질분임조)여부
HRM	조직직무 설계	직급단순화 실시 여부
	인적자원 계획	인성면접 여부, 집단토론 여부, HR업무담당전담조직 유무, 정기적인인력계획수립 여부
	정보체계	직무관리시스템 여부, 인사정보시스템 여부
	보상 및 장려	호봉제 여부, 연봉제 여부, 직무급 여부, 직능급 여부, 이익배분제도 여부, 전사성과급 여부
	복리후생	사원1년차 복리후생수준
	노사관계	노사관계
	기업규모	기업규모

표 4.2: 최적조합 선택법에서 변수 개수에 따른 카이제곱 값

변수개수(<i>i</i>)	카이제곱 값	<i>i</i> 번째와 (<i>i</i> - 1) 번째 카이제곱 값 차이
20	83.1038	
21	84.4531	1.3493
22	85.9601	1.5070
23	87.3558	1.3957
24	88.4126	1.0568
25	89.3819	0.9693
26	90.4624	1.0805
27	91.5984	1.1360
28	92.9092	1.3108

표 4.2는 단계적 선택법 및 후진소거법에 의해 선택된 변수 28개를 이용하여 최적조합 선택법으로 변수를 선택할 경우, 변수 개수가 각각 20~28개 일 때의 카이제곱(Chi-square)값을 나타낸 것이다. 20개의 변수 리스트부터 시작하여 하나의 변수를 추가할 때마다 모형 적합도의 값(score value)인 카이제곱 값의 변화를 관찰하면서 이 값이 가장 큰 차이가 날 때의 변수 개수를 선택하였다. 그 결과 최종적으로 선택된 변수는 22개였다.

선택된 변수의 최종 결과는 표 4.3에 나와 있다. 영역별로 세분화해서 살펴보면 대분류

표 4.3: 최적조합 선택법에 의해 선택된 변수

대분류	중분류	소분류(설문문항)
HRD	개인개발	국내연수 여부, 해외연수 여부
	경력개발	사내공모제 여부, 국가자격수당제도 여부
	조직개발	HR기능-성과및역량평가, QC(품질분임조)
HRM	조직직무 설계	직급단순화 실시 여부
	인적자원 계획	인성면접, 집단토론, HR업무담당전담조직유무, 정기적인인력계획수립
	정보체계	직무관리시스템, 인사정보시스템
	보상 및 장려	호봉제여부, 연봉제여부, 직무급여부, 직능급여부, 이익배분제도여부, 전사성과급여부
	복지후생	사원1년차복리후생수준
	노사관계	노사관계
기타	기업규모	기업규모

기준으로 HRD 영역 6개, HRM 영역 15개, 기타 1개이고, 중분류 기준으로 개인개발 2개, 경력개발 2개, 조직개발 2개, 조직직무설계 1개, 인적자원계획 4개, 정보체계 2개, 보상 및 장려 6개, 복지후생 1개, 노사관계 1개, 기업규모 1개가 선택되었다.

다음 절에서는 위의 표 4.3의 변수들을 이용하여 최종 로지스틱 회귀모형을 구축하고, 이렇게 구해진 모형의 결과값을 이용해 모형을 평가하였다.

4.2. 최종 모형 및 평가

기업 인적자원 관련 변수를 이용한 기업 신용점수 모형을 구축하기 위한 종속변수의 분포는 표 4.4와 같다. 한국신용평가(주) 기업 신용점수와 한국직업능력개발원의 설문문을 통합한 결과 종속변수인 기업 신용점수가 있는 100인 이상의 종업원을 가진 기업체수는 412개였다. 이 중에서 종속변수인 기업신용평가 점수는 70점 이상과 70점 미만이 각각 54.13%, 45.87%를 차지하고 있다. 종속변수의 두 범주의 비율이 큰 차이가 없이 균등에 가깝기 때문에 오분류율을 구할 때 분류절단값(cut-off value)은 0.5를 기준으로 하였다.

표 4.4: 종속변수 분포

(단위: 개, %)

구분	70점 미만	70점 이상
범주별 회사수	223(54.13)	189(45.87)

독립변수 중 기업규모는 총 4개의 범주로 구성되어 있으며 분포는 100~299명 41.99%, 300~999명 42.23%, 1000~1999명 8.01%, 2000명 이상이 7.77%를 차지하고 있다. 모형 구축에 사용된 기업규모의 약 84%가 1000명 미만의 기업으로부터 추출되었음을 보여주고 있으며, 나머지 독립변수들의 분포는 그림 4.1과 그림 4.2를 참고하기 바란다.

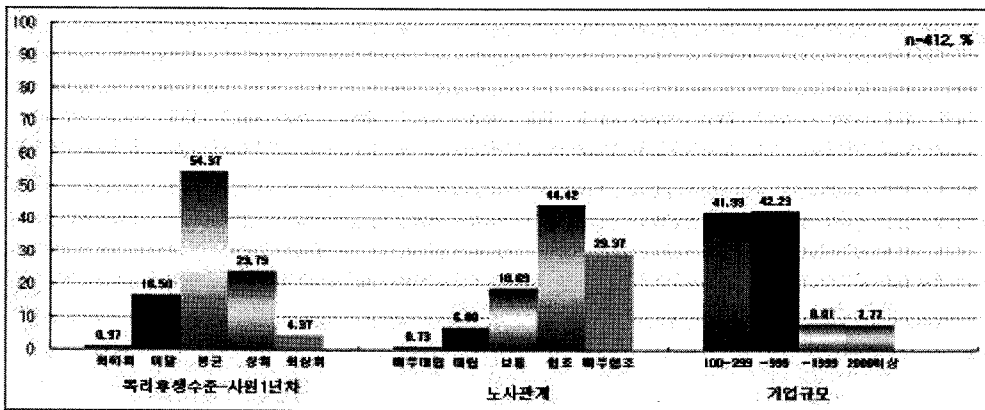


그림 4.1: 순위형 독립변수에 대한 분포

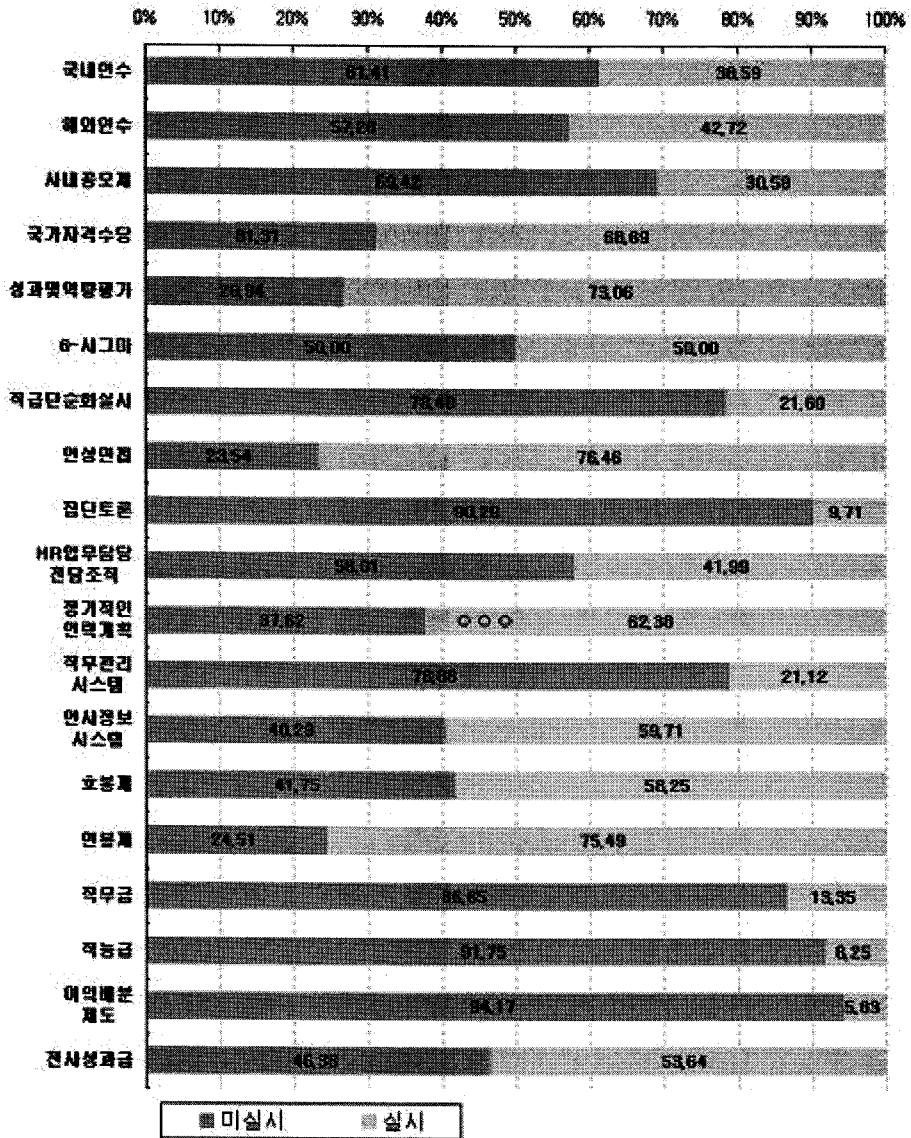


그림 4.2: 이진형 독립변수에 대한 분포

위의 22개 변수들을 이용하여 최종 모형을 구축한 결과가 표 4.5에 나와 있다. 표 4.5의 결과를 이용한 기업 인적자원 관련 활동에 대한 스코어 산출 모형은 다음의 식과 같다.

$$\begin{aligned}
\log(p/(1-p)) = & -2.3137 + 0.4517 * \text{국내연수여부} - 0.4062 * \text{해외연수여부} \\
& -0.4229 * \text{사내공모제여부} + 0.4972 * \text{국가자격수당여부} \\
& +0.5220 * \text{성과및역량평가여부} - 0.7327 * \text{품질분임조여부} \\
& +0.4876 * \text{직급단순화실시여부} - 0.4785 * \text{인성면접여부} \\
& -0.7254 * \text{집단토론여부} + 0.7803 * \text{HR업무담당전담조직유무} \\
& +0.4203 * \text{정기적인인력계획수립여부} - 0.6931 * \text{직무관리시스템여부} \\
& -0.3832 * \text{인사정보시스템여부} - 0.4971 * \text{호봉제여부} \\
& -0.8424 * \text{연봉제여부} - 0.6610 * \text{직무급여부} + 0.9349 * \text{직능급여부} \\
& +1.3566 * \text{이익배분제도여부} + 0.9108 * \text{전사성과급여부} \\
& +0.4126 * \text{사원1년차복리후생수준} + 0.1787 * \text{노사관계} \\
& +0.1989 * \text{기업규모}
\end{aligned} \tag{4.1}$$

여기에서 p 는 $\Pr(Y = 1 | x'_i s)$, 즉, 독립변수들이 주어졌을 때 기업 신용점수가 70점 이상($Y = 1$)일 확률을 의미한다.

표 4.5에서 오즈비는 독립변수의 값이 한단위 증가할 때 오즈 $p(1-p)$ 의 비에 대한 추정치이다. 오즈비가 1이라는 것은 $x'_i s$ 와 Y 가 독립이고, 1에서 멀어질수록 $x'_i s$ 가 Y 에 많은 영향을 미친다는 것을 의미한다.

그러므로 기업 신용점수가 70점 이상이 될 가능성(오즈)에 증가하는 방향으로 영향을 주는 항목은 국내연수여부, 국가자격수당제도여부, HR기능-성과및역량평가여부, 직급단순화실시여부, HR업무담당전담조직유무, 정기적인인력계획수립여부, 직능급여부, 이익배분제도여부, 전사성과급여부, 사원1년차복리후생수준, 노사관계, 기업규모였다. 특히, HR업무담당전담조직유무, 직능급여부, 이익배분제도여부, 전사성과급여부는 오즈비가 각각 2.2, 2.5, 3.9, 2.5로써, 이러한 항목들을 실시하는 기업체들이 실시하지 않는 기업체보다 신용점수가 70점 이상일 가능성이 약 2~4배였다. 이는 이러한 항목들을 실시하는 기업체들이 실시하지 않는 기업체들 보다 신용점수가 높음을 의미하므로, 각 기업체는 HR업무담당전담조직, 직능급, 이익배분제도, 전사성과급 등의 제도를 실시하는 것이 좋음을 의미한다.

기업 신용점수가 70점 이상이 될 가능성에 감소하는 방향으로 영향을 주는 항목은 해외연수여부, 사내공모제여부, QC(품질분임조)여부, 인성면접여부, 집단토론여부, 직무관리시스템여부, 인사정보시스템여부, 호봉제여부, 연봉제여부, 직무급여부였다. 특히, QC(품질분임조)여부, 집단토론여부, 연봉제여부, 직무급여부는 오즈비가 각각 0.48, 0.48, 0.43, 0.52로써, 이러한 항목들을 실시하지 않는 기업체가 실시하는 기업체보다 신용점수가 70점 이상일 가능성이 약 2.1~2.3배 정도 높았다. 이러한 항목들을 실시하지 않는 것이 기업의 신용점수를 높이는데 도움이 됨을 의미한다.

결론적으로 HR업무담당전담조직, 직능급, 이익배분제도, 전사성과급을 실시하고, QC(품질분임조), 채용 시 집단토론, 연봉제, 직무급을 실시하지 않는 것이 기업의 신용점수를

표 4.5: 최종 구축 모형 결과

대분류	중분류	소분류(설문)	추정치	p-값	오즈비
		절편	-2.3137	0.0062	
HRD	개인개발	국내연수 여부	0.4517	0.0870	1.571
		해외연수 여부	-0.4062	0.1320	0.666
	경력개발	사내공모제 여부	-0.4229	0.1040	0.655
		국가자격수당제도 여부	0.4972	0.0594	1.644
	조직개발	HR기능- 성과 및 역량평가 여부	0.5220	0.0739	1.685
		QC(품질분임조)여부	-0.7327	0.0030	0.481
HRM	조직직무 설계	직급단순화실시 여부	0.4876	0.0881	1.628
	인적자원 계획	인성면접 여부	-0.4785	0.0753	0.620
		집단토론 여부	-0.7254	0.0810	0.484
		HR업무담당전담조직 유무	0.7803	0.0056	2.182
		정기적인인력계획수립 여부	0.4203	0.1092	1.522
	정보체계	직무관리시스템 여부	-0.6931	0.0339	0.500
		인사정보시스템 여부	-0.3832	0.1606	0.682
	보상 및 장려	호봉제 여부	-0.4971	0.0499	0.608
		연봉제 여부	-0.8424	0.0062	0.431
		직무급 여부	-0.6610	0.0971	0.516
		직능급 여부	0.9349	0.0517	2.547
		이익배분제도 여부	1.3566	0.0096	3.883
		전사성과급 여부	0.9108	0.0003	2.486
복지후생	사원1년차복리후생수준	0.4126	0.0106	1.511	
노사관계	노사관계	0.1787	0.1710	1.196	
기타	기업규모	기업규모	0.1989	0.2080	1.220

높이는데 좋음을 알 수 있다. 오즈비를 살펴보았을 때 인적자원개발(HRD) 보다는 인적자원관리(HRM) 관련 항목들에 많은 투자나 관심을 기울여야 함을 알 수 있다. 특히, 인적자원관리 항목 중 보상 및 장려 영역이 기업 신용점수를 높이는데 많은 기여를 하므로 이 부분에 기업체는 많은 신경을 써야한다.

모형 구축에 사용된 기업체수가 412개로써 작기 때문에 412개 모든 기업체를 이용하여 로지스틱 회귀모형을 구축하고 그것을 평가하기 위해 10등급 교차타당성 분석결과가 표 4.6, 표 4.7, 그림 4.3에 나와 있다. 결과를 살펴보면 오분류율, G-mean은 각각 30.81, 68.27로써 나쁘지 않음을 알 수 있다. 그리고 반응율은 가장 좋은 십분위(decile = 0)는 가장 좋지 않은 십분위(decile = 9)보다 반응율이 약 6.08배가 크고 중간에 증가하는 부분이 있지만 대

표 4.6: 오분류표

구분	값
오분류율	30.81
G-mean	68.27

표 4.7: 반응율표

Decile	예측된 확률	반응율
0	0.8447	0.9083
1	0.7389	0.6917
2	0.6555	0.6167
3	0.5665	0.5917
4	0.5041	0.4683
5	0.4303	0.4971
6	0.3458	0.2333
7	0.2741	0.2750
8	0.2102	0.2033
9	0.1260	0.1493
Total	0.4595	0.4560

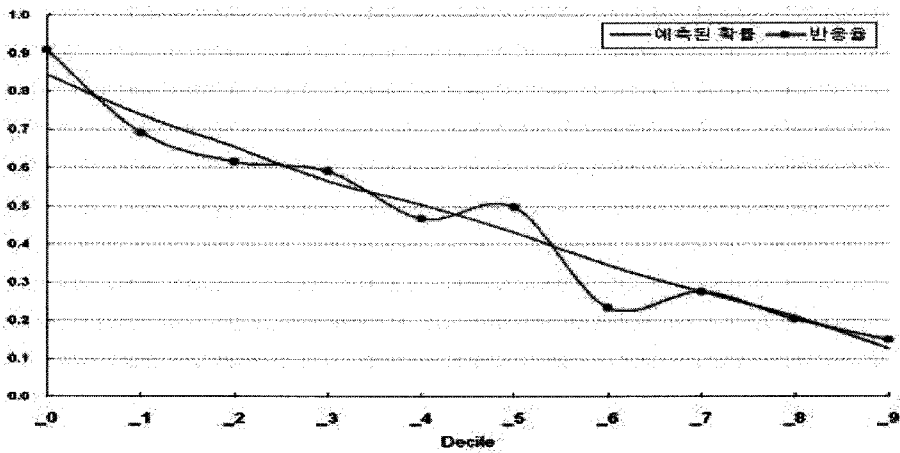


그림 4.3: 반응율 도표

제로 점차 감소는 경향을 보이고 있다. 그러므로 구축된 모형은 기업 인적자원 요소들을 이용하여 기업 신용점수를 측정하는데 적당한 모형이라는 결론을 내릴 수 있다.

5. 결론 및 향후 과제

본 논문의 목적은 신용점수에 영향을 미치는 기업 인적자원 요소들을 찾아서 신용점수 모형을 구축하는 것이다. 이를 위해 로지스틱 회귀모형을 구축하여 10등급교차타당성 분석을 통해 모형을 비교 및 평가하였다.

모형 구축 결과 최종적으로 선택된 변수는 22개였다. 영역별로는 대분류 기준으로 HRD 영역 6개, HRM 영역 15개, 기타 1개이고, 중분류 기준으로 개인개발 2개, 경력개발 2개, 조직개발 2개, 조직직무설계 1개, 인적자원계획 4개, 정보체계 2개, 보상 및 장려 6개, 복지후생 1개, 노사관계 1개, 기업규모 1개가 선택되었다.

최종적으로 선택된 변수는 국내연수여부, 해외연수여부, 사내공모제여부, 국가자격수당제도여부, HR기능(성과 및 역량평가)여부, QC(품질분임조)여부, 직급단순화실시여부, 인성면접여부, 집단토론여부, HR업무담당전담조직유무, 정기적인인력계획수립여부, 직무관리시스템여부, 인사정보시스템여부, 호봉제여부, 연봉제여부, 직무급여부, 직능급여부, 이익배분제도여부, 전사성과급여부, 사원1년차복리후생수준, 노사관계, 기업규모였다.

기업 신용점수가 70점 이상이 될 가능성에 증가하는 방향으로 영향을 주는 항목들은 국내연수여부, 국가자격수당제도여부, HR기능-성과및역량평가여부, 직급단순화실시여부, HR업무담당전담조직유무, 정기적인인력계획수립여부, 직능급여부, 이익배분제도여부, 전사성과급여부, 사원1년차복리후생수준, 노사관계, 기업규모였다. 특히, HR업무담당전담조직유무, 직능급여부, 이익배분제도여부, 전사성과급여부는 오즈비가 각각 2.2, 2.5, 3.9, 2.5로써, 이 항목들을 실시하는 기업체는 실시하지 않는 곳보다 기업 신용점수가 70점 이상일 가능성이 약 2~4배 높았다. 이는 이러한 항목들을 실시하는 기업체들이 실시하지 않는 기업체들 보다 신용점수가 높음을 의미하므로, 각 기업체는 HR업무담당전담조직, 직능급, 이익배분제도, 전사성과급 등의 제도를 실시하는 것이 좋음을 의미한다.

기업 신용점수가 70점 이상이 될 가능성에 감소하는 방향으로 영향을 주는 항목은 해외연수여부, 사내공모제여부, QC(품질분임조)여부, 인성면접여부, 집단토론여부, 직무관리시스템여부, 인사정보시스템여부, 호봉제여부, 연봉제여부, 직무급여부였다. 특히 QC(품질분임조)여부, 집단토론여부, 연봉제여부, 직무급여부는 오즈비가 각각 0.48, 0.48, 0.43, 0.52로써, 이러한 항목들을 실시하지 않는 기업체가 실시하는 기업체보다 신용점수가 70점 이상일 가능성이 약 2.1~2.3배 정도 높았다. 이러한 항목들을 실시하지 않는 것이 기업의 신용점수를 높이는데 도움이 됨을 의미한다.

결론적으로 HR업무담당전담조직, 직능급, 이익배분제도, 전사성과급은 실시하고, QC(품질분임조), 채용 시 집단토론, 연봉제, 직무급은 실시하지 않는 것이 기업의 신용점수를 높이는데 좋음을 알 수 있다. 기업의 신용점수를 높이기 위해서는 인적자원개발(HRD) 보다는 인적자원관리(HRM) 관련 항목들에 많은 투자나 관심을 기울여야 하는 것임을 알 수 있다. 특히, 기업 신용점수를 높이기 위해선 인적자원관리 항목 중 보상 및 장려에 많은 신경

을 써야한다.

모형 구축에 사용된 기업체수가 412개로써 작으므로 모형 평가를 위해 10등급 교차타당성 분석 결과 오분류를, G -mean은 각각 30.81, 68.27로써 나쁘지 않았다. 그리고 반응율은 가장 좋은 십분위는 가장 좋지 않은 십분위보다 약 6.08배 크고 점차 감소는 경향을 보이고 있다. 그러므로 구축된 모형은 기업 인적자원 요소들을 이용하여 기업 신용점수를 측정하는데 적당한 모형이라는 결론을 내릴 수 있다.

본 논문의 의의는 데이터마이닝 기법을 이용하여 기업 인적자원 요소들을 이용한 신용점수 모형 구축을 통해 인적자원 활동의 간접적인 경영 성과 기여도 측정을 위한 최초의 시도라는 점이다. 하지만 여러 가지 측면에서 한계를 내포하고 있다.

본 논문이 가지는 한계점과 향후 연구 방향은 다음과 같다. 첫째, 통계적 방법론 및 데이터마이닝 기법을 이용한 기업 인적자원 활동을 측정하기 위한 모형 구축 사례가 없어, 여기서는 인적자원 활동의 비용 효과성 측정 활동을 중심으로 모형을 구축하였다. 둘째, 본 논문에서 사용된 인적자원관련 활동 변수들은 일반화를 시키기에는 아직 부족한 부분이 많다. 그리고 투입된 변수들 역시 각 영역별로 정확히 부합되었는지에 대해서도 고려해야 할 점이 많다. 이는 주어진 자료를 연구 목적에 맞도록 인위적으로 결합하는 데서 드러난 한계에서 출발하기 때문이다. 따라서 향후의 연구를 위해서는 독립변수에 대한 보다 치밀한 논의가 필요하다. 신뢰성 검증 및 설문요인(factor)에 대한 탐색적 분석 및 확증적요인분석 등을 통해 타당성 검증을 하여야 할 필요가 있다. 셋째, 기업의 규모 및 업종에 따른 모형의 구축을 고려해보아야 한다. 기업의 규모와 업종에 따른 인적자원개발 및 관리 기법, 기업의 문화에 차이가 있을 수 있으므로 규모와 업종에 따라 어떠한 인적자원관련 항목들이 영향을 주는지 부가적인 연구가 필요하다. 마지막으로 기업 인적자원 요소들이 기업 신용도에 영향이 있음을 다양한 방법을 통해 추가적으로 규명한 후 기업 신용점수에 인적자원 요소들 중 일부 중요 요소를 포함시켜 새로운 형태의 신용점수 모형 구축에 대해서도 논의해 볼 필요가 있다.

6. 감사의 글

본 연구를 위해 자료 제공과 실증 분석을 위해 도움을 주신 한국직업능력개발원과 송창용박사님께 감사드립니다.

참고문헌

- 강현철, 한상태, 최종후, 김은석, 김미경 (1999). <SAS Enterprise Miner를 이용한 데이터마이닝-방법론 및 활용->, 자유아카데미.
- 강혜영 (2003). <인적자원관리에서의 ROI 분석 : 개념과 시사점, 포스코 인사실 포괄과제>, 포스코.
- 김미숙, 김안국, 이기성, 김재구, 이석재, 김태준 (2005). <인적자원개발 우수기관 인증제도 도입을 위한 심사지표 및 메뉴얼 개발 연구>, 한국직업능력개발원.

- 김용민 (2006). 기업 HCI 구성과 효과 측정을 위한 제언, Fellow Forum, 한국직업능력개발원.
- 송창용, 이성, 박주완 (2006a). 기업 HR 관련 활동과 기업 신용수준의 상관성 연구, 제1회 인적자본 기업패널 학술대회 논문집.
- 송창용, 이성, 박주완 (2006b). 중소기업의 기업 HR 활동 수준과 신용수준의 상관성 연구, <직업능력개발연구>, 9, 171-192.
- 이성 (2003). <혁신적 HR 성공전략 ROI>, 학지사
- 이영섭 (2003). <데이터마이닝 Cookbook>, 교우사
- 한국신용평가 (2005). 정보화 수준과 기업신용평가 수준과의 연관성 연구에 대한 자문내역보고서.
- Becker B. E., Huselid, M. A. and Ulrich, D. (2001). *The HR Scorecard: Linking People, Strategy, and Performance*, Harvard Business School Press.
- Hosmer, D. W. and Lemeshow, S. (2000). *Applied Logistic Regression*, 2nd ed., John Wiley & Sons, New York.
- Huselid, M. A., Becker, B. E. and Beatty, R. W. (2005). *The Workforce Scorecard: Managing Human Capital to Execute Strategy*, Harvard Business School Press.
- McLagan, P. (1989). *Models of HRD Practice*, ASTD Press.

[2007년 3월 접수, 2007년 4월 채택]

A Study for Building Credit Scoring Model using Enterprise Human Resource Factors*

Yung-Seop Lee¹⁾ Joo-Wan Park²⁾

ABSTRACT

Although various models have been developed to establish the enterprise credit scoring, no model has utilized the enterprise human resource so far. The purpose of this study was to build an enterprise credit scoring model using enterprise human resource factors.

The data to measure the enterprise credit score were made by the first-year research material of HCCP was used to investigate the enterprise human resource and 2004 Credit Rating Score generated from KIS-Credit Scoring Model. The independent variables were chosen among questionnaires of HCCP based on Mclagan(1989)'s HR wheel model, and the credit score of Korean Information Service was used for the dependent variables. The statistical method used for data analysis was logistic regression.

As a result of constructing a model, 22 variables were selected. To see these specifically by each large area, 6 variables in human resource development(HRD) area, 15 in human resource management(HRM) area, and 1 in the other area were chosen.

As a consequence of 10 fold cross validation, misclassification rate and G-mean were 30.81 and 68.27 respectively. Decile having the highest response rate was bigger than the one having the lowest response rate by 6.08 times, and had a tendency to decrease. Therefore, the result of study showed that the proposed model was appropriate to measure enterprise credit score using enterprise human resource variables.

Keywords: Data mining, logistic regression, credit score, Human Resource Management (HRM), Human Resource Development (HRD).

* This work was supported by grant No. R01-2004-10689-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

1) Associate professor, Department of Statistics, Dongguk University, Pil-dong 2ga, Jung-gu, Seoul 100-715, Korea
E-mail: yung@dongguk.edu

2) Graduate Student, Department of Statistics, Dongguk University, Pil-dong 2ga, Jung-gu, Seoul 100-715, Korea.
E-mail: jwan0217@naver.com