

▣ 연구논문

Knowledge Discovery in Databases에 기반한 경영성과분석 시스템

Business Performance Analysis System based on Knowledge Discovery in Databases

조성훈

Seong-Hoon Cho*

정민용

Min-Yong Chung**

Abstract

In dynamic management environment, CEO must make an efficient decision with information & knowledge management systems based on IT(Information Technology). As a key component to cope with this current, we suggest the business performance analysis system based on KDD(Knowledge Discovery in Databases).

We consider the theoretical model that is composited both Value-Added in respect of stakeholder and Economic Value-Added in respect of shareholder. Additionally we use DBMS and data mining method using Genetic Algorithms as physical model.

To demonstrate the performance of the business performance analysis system, we analyse a domestic motors industry. The empirical case is based on the financial data of KISFAS(Korea Investors Services Financial Analysis System) database. The samples included in the study consist of H motors/S motors industry over the 16-year from 1981 to 1996.

1. 서론

기업의 최고이사결정자인 CEO는 급속하게 변화하는 경영환경에 대응하기 위해서, 기업의 전 조직이 정보·지식을 동시에 공유할 수 있는 효율적인 의사결정 체계를 요구하고 있다. 이러한 의미에서 기업은 정보기술을 이용하여 적절하게 통합된 정보·지식을 꼭 필요로 하는 부서에 적시에 제공할 수 있어야 하며, 정보·지식 공유를 통하여 모든 조직원의 역량을 기업 목표라는 초점에 집중시킬 수 있는 기반을 갖고 있어야 한다.

따라서 기업의 경영관리영역인 재무(Finance)·생산(Manufacturing)·판매(Marketing) 등에서 발생하는 기업의 현황을 자동화된 정보체계로 흡수·축적하여 조직 전체가 공유할 수 있는 정보·지식으로 변환하는 체계를 구축하는 것이 기업의 생존에 중요한 요소로 부각되고 있다.

이에 본 연구는 이러한 기업경영환경에 부응하기 위하여 경영성과지표를 관리하는 의사결정자의 관점에서 정보기술의 효과적인 활용을 통하여 기업 전체의 조직이 공유할 수 있는 경영성과분석 체계를 전개하고자 한다. 이를 위하여 기업의 경영성과분석에 대한 일반론을 서술

* 삼성전자 경영혁신팀

** 건국대학교 산업공학과

하여 이해관계자 관점에서의 부가가치(Value-Added)와 자본투자가 관점에서의 경제적부가가치(Economic Value-Added)를 기업 경영성과의 두 축으로 제시한다.

그리고 KDD(Knowledge Discovery in Databases)라는 체계속에서 경영성과에 민감하게 반응하는 중요관리변수를 추출하기 위하여 데이터베이스(Database)내의 지식발견(Knowledge Discovery)에 기반한 데이터마이닝(Data Mining)시스템을 구현한다.

본 연구에서 제시하는 체계는 경영성과지표를 관리하는 의사결정자의 관점에서, 정보기술의 효과적인 활용을 통하여 기업 전체의 조직이 공유할 수 있는 경영성과분석체계의 바탕이 될 것이라 판단한다.

2. 기업관리체계에서의 경영성과분석

현대의 기업경영관리는 부가가치 또는 시장점유율의 극대화를 추구하는 일본·독일형 기업과 이윤극대화를 강조하는 미국형이 대조를 이루고 있다. 우리나라의 경우는 일반적으로 미국형의 수익을 강조하는 경영정책 및 원칙이 두드러져 보이나 장·단기적으로 기업의 유지·성장을 도모하기 위해서는 장기적 부가가치창출력과 단기적 이익을 동시에 고려할 수 있는 기업관리체계가 절실히 요청된다고 판단된다.

2.1. 부가가치(Value-Added)를 중심으로 한 경영성과분석

오늘날 기업에서는 기업의 성과를 부가가치로 파악하고 그 중에서도 부가가치생산성이란 개념으로 일반화하고 있는데, 이는 기업의 경영 합리화를 통하여 고임금·고능률·고이윤 등의 목표를 실현하고, 종래의 명목적 이익인 매출액 중심의 경영에서 실질적 이익이라 할 수 있는 부가가치 중심의 경영으로 전환하려 하기 때문이다.

부가가치를 산정하는 목적은 다음과 같다.

- 첫째, 기업의 상대적 평가수단으로 이용된다. 즉 타 기업의 부가가치생산성과 비교하여 어느 정도 합리성을 가지고 있는가 혹은 기간비교에 있어서 어느 정도 부가가치가 증가되고 있는가 등을 비교함으로써 상대적인 평가가 가능해진다.
- 둘째, 기업 업적의 동태적 분석의 수단으로 이용된다. 즉 투입된 노동량이나 자본량과 같은 기업 투자가치의 증감이 기업의 산출 가치량에 어떠한 영향을 미치는가 등을 파악할 수 있다.
- 셋째, 기업의 사회적 공헌도를 표시하는 공적으로서, 획득한 부가가치가 노동-자본-사회로 구분되는 기업 이해관계자에게 어떻게 분배되고 있는가를 나타내기 위하여 사용된다.

일반적 의미에서의 부가가치생산성은 노동력에 의한 부가가치의 획득을 의미하지만, 실제로 부가가치가 단지 노동력에 의해서만 창출되는 것은 아니며, 자본력도 부가가치 창출에 크게 기여하고 있다는 사실을 간과해서는 안된다. 특히 생산체제는 하나의 투입요소로서 이루어지는 것이 아니라 전체 요소의 결합으로서 이루어진다는 점을 감안한다면, 부가가치를 노동, 자본 그리고 기타 투입물에 의한 성과로 간주해야 한다는 일반론은 타당성이 있다고 판단된다.

이러한 부가가치생산성은 기업의 이윤 증대에 대한 근로자 기여분의 공정한 분배를 현실적으로 파악할 수 있는 좋은 기준이 될 수 있으므로, 기업의 사회적 공헌이 중시되고 있는 추세를 적절히 반영한다고 할 수 있다.

2.2. 경제적부가가치(Economic Value-Added)를 중심으로 한 경영성과분석

세계경쟁이 치열해지고 불확실성이 증폭되는, 고비용·저효율이라는 새로운 경제환경하에서는 기업의 경영패러다임이 외형중심에서 가치중심경영(Value Based Management)으로 전환되어야 한다. 경제적부가가치를 중심으로 한 가치중심경영은 기업경영의 최우선순위를 가치창출에 두는 것을 의미하며, 현금흐름(Cash Flow)을 장기적으로 최대화하는 것을 목표로 한다. 가치중심경영은 Coca-Cola, AT&T와 같은 외국기업은 물론, 포항제철, LG 등 몇몇 국내 기업에서도 이미 도입되었다. 따라서 가치중심경영은 국제적 표준(Global Standard)으로 정착되어 가는 과정에 있다고 할 수 있다.

경제적부가가치는 미국의 컨설팅회사 Stern Stewart사가 1980후반에 도입한 개념으로서, 기업이 영업활동을 통해 창출한 순가치의 증가분인 세후순영업이익에서 투입된 자본에 대한 자본코스트를 차감한 금액을 나타낸 것이다.[4] 과거에 중시되던 당기순이익을 기준으로 한 경영지표는 타인자본에 대한 비용만을 반영하고 있으므로 자기자본비용을 반영하지 못하였으나, 이에 반해 경제적부가가치는 기회 비용이 반영된 주주의 기대수익을 자본비용에 포함시킴으로서 주주의 기회비용을 명확히 할 수 있다.

또한 경제적부가가치는 기업이 영업활동을 통해 달성한 이익에서 이를 위해 투입한 자본에 대한 정당한 자본비용을 차감한 후에 성취된 이익의 양을 측정하는 것이다. 이러한 의미에서 경제적부가가치는 기업의 근본적인 생산활동만을 그 대상으로 삼고 있다고 할 수 있다.

결국 경제적 부가가치를 통하여 기업을 관리한다는 것은 경영활동에 투입된 자본을 중점적으로 강조하여 자기자본을 포함한 모든 자본비용을 감안하고도 여분의 수익을 성취해야만, 기업이 진정한 의미의 이익을 창출하여 주주의 부를 극대화시킬 수 있다는 것을 의미한다고 할 수 있다.[10]

2.3. 부가가치와 경제적부가가치를 결합한 경영성과분석 구조

전술한 바와 같이 시대적 흐름이나 지역적 차이에 따라 부가가치와 이익, 최근에는 이익을 대신하여 경제적부가가치가 혼재하여 경영성과의 측정기준으로서 사용되고 있음을 알 수 있다. 전술한 두 가지 목표는 각국의 시장경제의 발달 및 사회·문화적 환경에 따라 크게 2가지 형태로 발달되어왔는데 하나는 미국·영국을 중심으로 자유경쟁의 원리하에서 주주의 부를 극대화하는 주주 자본주의와 다른 하나는 독일·일본을 중심으로 국가의 주도아래 기업을 둘러싼 여러 이해관계자의 이익을 중시하는 이해관계자 자본주의이다. 이를 정리하면 다음의 <표 1>과 같다.[5]

<표 1> 이해관계자 자본주의와 주주 자본주의의 비교

	이해관계자 자본주의(Stakeholder Capitalism)	주주 자본주의(Shareholder Capitalism)
대표국가	독일, 일본	미국, 영국
경영목표	기업에 관계된 모든 계층에 이익제공 -은행중시(채권자와 장기적인 협력) -장기전략중시(시장점유율)	주주에 대한 배당이익의 극대화 -단기수익성 -증권시장중시(증권시장 흐름 중시, 시가주의)
소유구조	종업원과 협력업체도 지분참여 -은행이나 계열사간의 상호주식보유	대주주와 배당이익목적 주식투자가들 중심 -주식분산으로 소유와 경영분리
임금체계	종신고용, 연공서열, 고정급	실적위주의 연봉제 중심
종업원신분	노조나 법적보호장치를 통해 신분 안전 확보	실적여부에 따라 승진과 실직기회 상존
복지재원	정부재정 외 기업부담 많음	정부재정 대부분 담당
대표회계	원가기획(기획원가)	ABC(Activity-Based Costing)

이러한 흐름에 대하여 김태성은 [부가가치/종업원수]와 [이익/자기자본]이라는 기업 재무평가의 두 가지 목표를 동시에 고찰하여야 한다는 생산성관리네트워크를 제시하였다. 이 연구에 의하면 한 기업의 경제 전반의 성장가능성을 측정하는데 중요한 요소로서 종업원 1인당 부가가치를 사용하며, 경영자층에서 일반적으로 다루어지는 경영성과분석의 척도로서 자본투자가의 이익과 관련하여 자기자본에 대한 이익(자기자본이익율)을 사용하였다.[3]

그러나 자기자본이익율에서 사용되는 이익은 회계적 이익개념으로서, 현행의 회계이익으로 경영성과를 평가하는 과정에서는 자기자본조달사용에 대한 대가로 지불되어야 할 자기자본비용을 비용으로 인식하지 않는다. 따라서 단순히 회계상의 법인세차감후순이익을 기준으로 경영성과를 평가하는 것은 자본제공자인 주주의 입장에서 판단한다면 자본조달에 따른 자본비용을 충분히 감안하지 못한다고 인식할 수 있으며, 사전에 예상된 자기자본비용 보상에 대한 주주의 중요관심사를 외면하게 된다.

회계적 이익이 기업의 경영성과를 올바르게 반영하지 못하는 이유는 다음과 같이 정리할 수 있다.

첫째, 기업의 경영성과가 좋을수록 회계이익은 증가한다. 그러나 회계이익이 큰 경우에도 주주의 요구수익률이 높은 경우 자기자본조달비용을 많이 지급해야하므로 총자본비용 보상후의 경영성과는 회계이익과 반드시 비례하지는 않는다. 실제로 기업의 경영상태가 불안한 경우에는 자기자본을 투하하는 주주는 위험률에 비례하여 많은 분량의 자기자본비용을 기대하게 된다.

둘째, 기업의 경영성과에서 이자비용을 공제하여 추정되는 회계이익이 크더라도 기업의 자기자본사용 규모가 큰 경우에는 주주에게 충분한 보상을 지급할 수 없는 경우가 발생할 수 있다. 이 경우 기업의 부채비율이 상대적으로 낮아 이자비용의 지급액은 작아지므로 기업의 회계이익은 커지지만, 자기자본사용부분에 대한 이자비용을 고려한다면 실질적인 기업의 이익은 줄어들 수 밖에 없다. 따라서 타인자본과 자기자본을 동시에 고려한다면 총자본비용 보상후의 경영성과는 회계이익과 반드시 비례관계에 있다고 할 수는 없다.

국내의 연구 결과, 1987년부터 1992년까지 EPS(Earning Per Share)가 모두 양의 값을 갖는 표본상장기업중 약 70~80%에 이르는 기업들의 「경제적부가가치/주(Share)」가 음(-)의 값을 갖는 것으로 나타났다.[2] 이는 상당수의 기업들이 영업활동을 통하여 얻는 수익성으로 자기자본조달비용을 충분히 보상하지 못하고 있음을 나타내는 결과라고 할 수 있다.[8]

결국 회계이익의 경우에는 타인자본비용만을 고정적으로 보상한 후의 경영성과이지만, 경제적부가가치는 타인자본비용과 자기자본비용을 모두 고려한 경영성과를 의미하기 때문에 자본제공자의 입장에서는 회계이익정보와 요구자기자본비용을 동시에 포함한 경제적부가가치의 개념이 더 유용한 정보원천으로 활용할 수 있다.

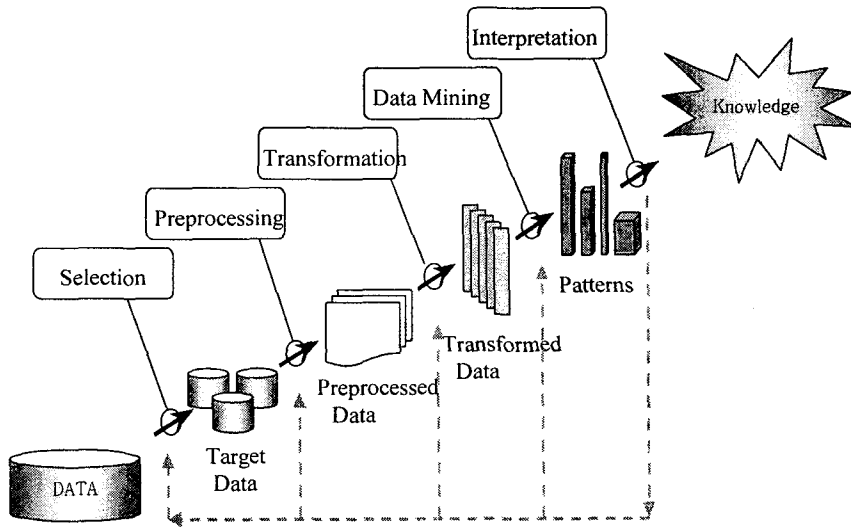
이와 같이 기업이 산출한 이익 개념으로서 회계상의 당기순이익보다는 영업관련 현금흐름에 입각한 경제적부가가치가 중시되는 최근의 경향을 중시하여, 본 연구에서는 기업의 경영성과를 평가하는 틀로서 이해관계자 입장에서의 1인당부가가치(부가가치/종업원수)와 자본제공자 입장에서의 투하자본대비경제적부가가치(EVA/평균투하자본)를 동시에 사용하고자 한다.

3. KDD에 기반한 경영성과분석 시스템의 이론모형

자료의 양이 기하급수적으로 증가하면서 자료를 정확히 분석하는 것이 점점 어려워지고 있다. 최근 들어 컴퓨터를 이용하여 자료를 처리하고, 막대한 양의 자료로부터 의미있는 변화의 경향과 구조를 탐색하게 되었다. 이와 같은 과정에서 KDD(Knowledge Discovery in

Databases)라는 체계가 대두되었다.

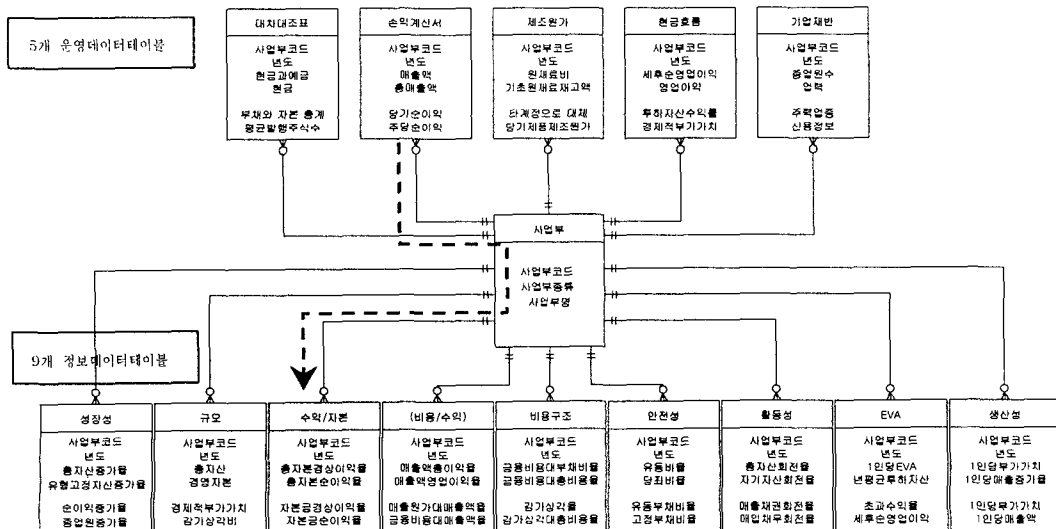
KDD의 전반적인 과정을 도시하면 다음의 <그림 1>과 같다. <그림 1>에서 제시된 바와 같이 KDD에서는 자료의 집합체인 데이터베이스(Database), 자료가공(Transformed Data), 데이터마이닝(Data Mining) 그리고 이에 대한 해석(Interpretation)까지 일련의 자동화된 과정을 포함하고 있음을 알 수 있다.[11]



<그림 1> Knowledge Discovery in Databases의 프로세스

3.1. 경영성과 데이터베이스

본 연구에서는 원시데이터베이스로서 기업의 재무제표를 사용하고, 사용자의 요구에 따라 원시데이터로부터 2개의 경영성과지표(1인당부가가치, 투자자본대비경제적부가가치)와 85개의 재무비율을 추출하여 데이터베이스에 저장한다.



<그림 2> 경영성과 데이터베이스의 개체-관계도

기업의 경영활동을 평가하기 위해서는 기업이 공시한 재무제표를 개별적으로 분리하여 판단할 수 없으며, 대차대조표, 손익계산서와 함께 제조원가명세서, 현금흐름 등도 동시에 고려해야 한다고 할 수 있다. 이에 본 연구에서는 기업경영의 결과로 발생하는 운영적 재무결과를 5개의 테이블로 분류하여, 「대차대조표」, 「손익계산서」, 「제조원가명세서」, 「현금흐름」, 「기업제반」 테이블을 구성하고 이를 기업의 경영성과분석을 위한 원시데이터로서 사용한다.

그리고 5개의 운영데이터(Operational Data) 테이블을 기반으로 「성장성」, 「규모」, 「수익성(수익/자본)」, 「수익성(수익/비용)」, 「수익성(비용구조)」, 「안정성」, 「활동성」, 「경제적부가가치」, 「생산성」의 9개 정보데이터(Information Data) 테이블을 생성·유지한다. 관리자이외의 분석목적으로 시스템에 접근하는 사용자들은 9개의 정보데이터 테이블에 접근하여 기업의 경영성과를 분석한다. 구축된 데이터베이스의 개체-관계도를 작성하면 다음의 <그림 2>와 같다.

<그림 2>에서 관리자는 서버를 통하여 원시데이터베이스로서 5개 운영데이터 테이블에 경영성과 자료를 입력하며, 클라이언트를 통하여 사용자의 요청이 발생하면 운영데이터 테이블을 기반으로 9개의 정보데이터 테이블이 생성된다. 실제로 클라이언트에서는 9개의 정보데이터 테이블에만 접근이 허용된다.

3.2. 유전알고리즘을 이용한 데이터마이닝

과거 자료내에 존재하는 변수들간의 관련성을 분석하는 도구로서 가장 많이 사용되었던 회귀분석(Regression Analysis), 인자분석(Factor Analysis), 구조방정식모형(Structural Equation Modeling) 등의 대부분의 통계적 분석기법은 선형성이나 정규성과 같은 가정에 바탕을 두었기 때문에, 매우 제한적으로 사용될 수 밖에 없다. 더욱이, 본 연구에서 분석의 대상으로 사용하는 경영성과 데이터베이스내의 재무관련 비율은 여러 요소가 중복적으로 비율의 분모와 분자로 사용되므로, 다중공선성의 문제가 자주 발생하기 때문에 통계적 분석기법을 사용하기에는 여러 가지 제약이 존재한다. 이러한 한계를 극복하고자 다양한 분야에서 인공지능망을 이용한 연구가 이루어져왔다. 기존의 연구결과를 통해서도 주식시장의 변화 예측, 도산기업 예측 등의 경영·재무분야에서 인공지능망은 기존의 복잡한 통계적 분석기법보다 더욱 좋은 결과를 보여왔다.[5]

그러나 인공지능망의 내부배열(은닉층(Hidden Layer)·은닉마디(Hidden Node))은 확실한 이론적 배경보다는 경우에 따라 시행착오 또는 어림짐작의 방법으로 결정되는 경우가 많으며, 이러한 이론적 토대의 부재가 인공지능망 모형의 응용연구에서 항상 걸림돌이 되어왔다. 또한 인공지능망 모형은 그 자체의 Black Box적 특성 때문에 입력변수와 출력변수간의 상호관계를 이해하기가 쉽지 않다. 통계모델에서 계수와 변수들의 확률값은 관계의 강도와 유의도를 나타내주는데 반하여, 인공지능망 모델에서 내부 가중치들은 입력층·은닉층·출력층내에서 모두 함께 얽혀 있어 변수들간의 상호관계를 분명하게 표출하지 못함으로 결과 도출의 이해력이 떨어진다. 또한 인공지능망의 가장 큰 단점으로 확인되는 것은 학습과정에서 해의 전역성을 확보하기가 어렵다는 것이다.[1][7][9]

이러한 한계를 감안하여 본 연구에서는 기계학습의 한 종류인 유전알고리즘을 통하여 경영성과의 중요 관리변수를 추출한다. 다양한 기계학습론중에서 유전알고리즘을 활용하는 이유는 통계적 분석기법과는 달리 선형성·정규성의 가정이 필요하지 않으며, 해공간이 넓고 자료의 노이즈가 심한 경우에도 비교적 전역해를 빠른 시간에 탐색할 수 있기 때문이다. 실제로 고전적인 통계기법은 데이터베이스와 같이 자료의 양이 방대한 경우 선형성이나 정규성을 만족하기가 더욱 어려우며, 이와 같은 측면에서 유전알고리즘과 같은 기계학습 기법이 유용하게 사용될 수 있다.

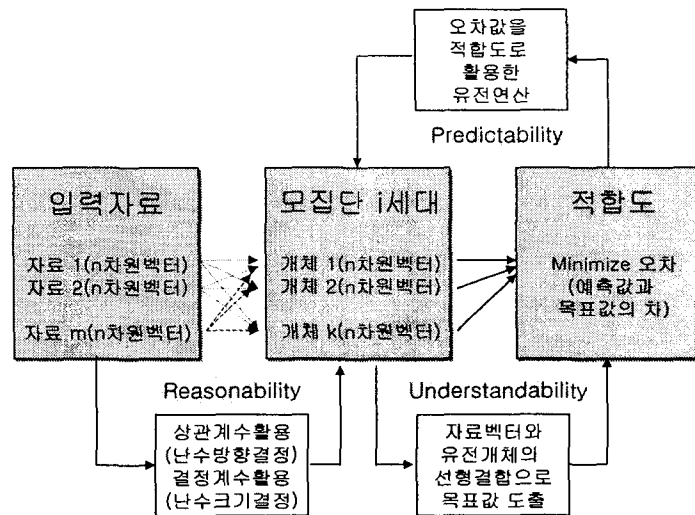
그러나 일반적인 유전알고리즘은 무작위로 발생된 난수값을 기반으로 평가·선택·교차·돌연변이 등의 연산자를 거치면서 추출된 귀납적 결과를 받아들여야하므로, 반/비 정형화된 문제의 경우에는 의사결정자가 그 결과를 전적으로 받아들이기가 쉽지 않다. 이에 본 연구에서는 선형결합의 구조하에 통계적 상관분석을 결합한 설명기반 유전알고리즘체계를 제시하고 이를 경영성과 데이터베이스의 데이터마이닝에 활용하고자 한다.

3.3. 설명기반 유전알고리즘

일반적인 유전알고리즘에서는 수식으로 표현되는 하나의 적합도 함수를 최적화하는 방식으로 최적해 산출의 기준을 설정하므로 구조화된 의사결정문제에는 적합할 수 있으나, 본 연구에서 지향하는 경영성과분석과 같이 다양한 변수간의 상호관계가 존재하며 의사결정자의 사전지식과 같이 비정형화된 요소들도 많이 존재하는 경우에는, 예측력(오차의 크기)과 같은 하나의 최적화함수만을 기준으로 최적해를 산출하고 이를 해석하는 것은 여러 가지 어려움이 존재할 수 있다.[4]

본 연구에서는 이러한 측면을 고려하여 유전알고리즘을 선형결합형태로 설계함으로써, 복잡한 비선형구조에 비하여 예측정도는 떨어질 수 있으나 입력변수와 결과변수간의 선형가중치를 변수의 중요도로 그대로 사용하여 도출된 지식의 이해력(결과의 의미)을 증진시킨다. 또한 유전알고리즘의 초기 유전자값을 일양분포 난수에 의하지 않고, 분석자료입력시 산출되는 입력변수와 결과변수의 상관계수(Correlation Coefficient)와 결정계수(Determination Coefficient)를 정규분포 난수의 Seed로 활용함으로써 추출된 지식의 타당성(결과의 수용도)을 높인다.

유전알고리즘을 활용한 중요 관리변수의 학습구조를 도시하면 다음의 <그림 3>과 같다.



<그림 3> 중요 관리변수 추출을 위한 유전알고리즘 학습구조

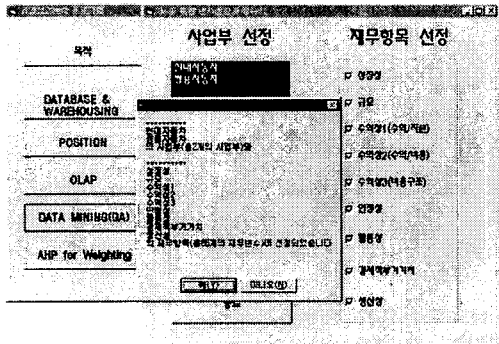
<그림 3>의 학습구조에서는 비선형모형으로서 예측력의 증대를 추구하지 않고 선형구조를 채택함으로써 추출된 중요 관리변수의 이해력을 도모하였으며, 정규분포를 활용한 초기 유전자값 설계과정을 통하여 지나친 탐색공간의 확대를 방지하고 또한 상관계수와 결정계수를 초기 유전값의 Seed로 사용함으로써 기계학습의 결과를 합리적으로 받아들일 수 있는 타당성측면을 고려하였다. 이와 같은 설명기반 유전알고리즘체계는 예측력이라는 요인이외에도 이해력·타당성이라는 특성을 동시에 추구한다고 할 수 있다.

4. 경영성과분석 시스템의 구축 및 활용

본 연구에서 구현한 경영성과분석 시스템은 관계형 데이터베이스를 기반으로 한 클라이언트-서버 구조에 바탕을 두고 있다. 서버의 운영시스템으로 윈도우 NT 4.0을 사용하고 데이터 웨어하우스는 MS-SQL 7.0을 사용해 원시 데이터베이스와 같은 시스템상에서 구현하였다. Mining서버는 윈도우 NT상에서 구현하며 MS-Visual Basic 6.0을 통하여 구축되었으며, Mining클라이언트는 윈도우 98상에서 구현하며 역시 MS-Visual Basic 6.0을 통하여 구현되었다.

4.1. KDD에 기반한 경영성과분석 시스템의 활용

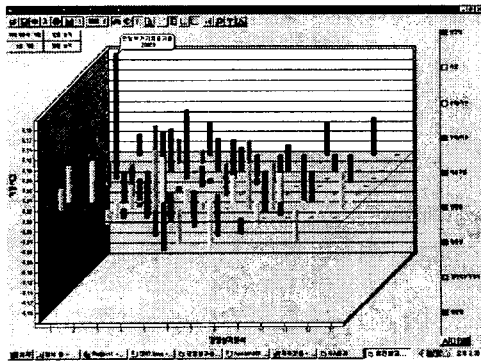
4.1.1. 데이터베이스내의 자료 추출



<그림 5> 데이터베이스내의 자료 추출

데이터마이닝을 위해서는 <그림 4>와 같이 사업부와 재무항목을 선정함으로써 분석자료를 추출하여야 한다. 여기서 선정된 자료는 유전알고리즘내의 학습자료로 활용되므로, <그림 4>와 같이 H자동차와 S자동차 모두를 선정하고 9개의 재무항목을 모두 선정하는 경우에는 85개의 변수(9개 재무항목내의 재무비율)를 포함하는 32개 학습자료가 생성된다. 이후에는 유전알고리즘에서 사용되는 목표변수의 가중치와 유전연산자의 모수값을 부여한다. 목표변수의 가중값은 0과 100사이로서, 1인당부가가치만으로 목표변수를 설정하여 1인당부가가치에 상대적으로 민감하게 변동하는 관리변수를 추출하고자 하는 경우에는 VA Weight를 100으로 설정한다.

4.1.2. 중요 관리변수의 시각화를 통한 해석



<그림 6> 중요 관리변수의 시각화

유전알고리즘의 수행 결과를 도시하면 <그림 6>과 같다. 최고가중치를 가지는 1인당부가가치 증가율의 경우 0.20658의 값이 가중치로서 산출되었으며, 가장 중요한 관리변수로 추출되었음을 알 수 있다.

4.2. 데이터마이닝체계의 유용성

전술한 바와 같이 본 연구에서 제시하는 유전알고리즘은 설명기반의 특성을 지향하기 위하여 예측력과 함께 타당성과 이해력을 동시에 고려한다. 이해력은 선형 학습구조로서 설명될 수 있으므로, 본 절에서는 타당성과 예측력에 대하여 서술한다.

4.2.1. 타당성 비교

통계적 상관분석의 특성을 적절히 반영하는가를 입증하기 위하여 <표 2>에서는 정규분포형 난수를 사용한 유전알고리즘과 일양분포형 난수를 사용한 유전알고리즘을 비교하였다. 물론

유전알고리즘의 타당성을 판단하는 기준으로 상관분석만을 사용할 수는 없지만, 전술한 바와 같이 통계적 상관분석은 전통적으로 변수간의 영향력을 파악하는 수단으로서 사용되었던 방법 이므로 설명기반이라는 관점에서 기계학습으로 추출된 지식의 타당성을 입증할 수 있다고 판단하였다.

본 연구에서 의도한 바와 같이 정규분포형 난수체계를 이용한 결과가 일양분포형 난수체계를 이용한 경우보다 상관분석의 결과에 근접하였음을 알 수 있으며, 특히 정규분포형 난수를 사용한 3000세대와 5000세대의 경우에는 상관계수가 5% 수준에서 유의한 것으로 나타났다.

<표 2> 상관분석을 기준으로 한 정규 난수와 일양 난수 유전알고리즘의 타당성 비교

비교항목 (세대수)	정규난수 (1000)	정규난수 (3000)	정규난수 (5000)	일양난수 (1000)	일양난수 (3000)	일양난수 (5000)
상관분석과의 적률상관	0.170	0.243*	0.232*	0.032	0.082	0.128

4.2.2. 예측력 비교

타당성과 함께 유전알고리즘 예측력을 비교하기 위하여 유전알고리즘의 오차와 비선형모형을 사용하는 인공신경망의 오차를 비교한다.

인공신경망기법으로서 후방학습알고리즘(Back Propagation Algorithm)을 적용하는 BrainMaker for Professional V3.11을 사용하였으며, 입력노드는 85개, 출력노드는 1개, 은닉층 1개 그리고 은닉노드 85개를 사용하였다. 32개의 학습자료를 모두 훈련용으로 사용하였으며, 5000회의 반복훈련중에서 가장 작은 RMS(Root Mean Squares)를 기준으로 인공신경망의 예측력을 평가한다.

<표 3>에서 알 수 있는 바와 같이, 비선형모형을 갖고 있는 인공신경망의 경우 오차를 나타내는 RMS가 유전알고리즘보다 큰 것을 알 수 있다. 이는 85개의 관리변수에 대한 85개 차원의 최적해 탐색이 수행됨으로서 전역해를 보장하지 못하는 인공신경망의 특성이 두드러지게 나타난 것이라 해석할 수 있다. 또한 정규분포형 난수를 사용한 유전알고리즘의 RMS가 일양분포형 난수를 사용한 유전알고리즘의 RMS보다 상대적으로 적은데, 이는 초기해의 범위를 적절하게 축소하여 해의 탐색에서 발생하는 낭비를 방지함으로써 나타난 결과로 해석할 수 있다.

<표 3> 정규 난수와 일양 난수 유전알고리즘의 예측력 비교

비교항목 (세대수)	신경망 (5000)	정규난수 (1000)	정규난수 (3000)	정규난수 (5000)	일양난수 (1000)	일양난수 (3000)	일양난수 (5000)
RMS	0.0367	0.0163	0.0154	0.0149	0.0163	0.0154	0.0149

5. 결론 및 향후연구

본 연구에서는 변화하는 기업환경에 부응하기 위하여, 경영성과지표를 관리하는 의사결정자의 관점에서 정보기술의 효과적인 활용을 통하여 기업 전체의 조직이 공유할 수 있는 경영성과분석 체계를 전개하였다. 이를 위하여 기업의 경영성과분석의 틀을 이해관계자 관점에서의 부가가치와 가치경영 관점에서의 경제적부가가치로서 제시하였으며, 제시된 경영성과분석 체계를 통하여 기업내부의 이해관계자는 물론 기업 외부의 이해관계자 역시 기업의 경영성과를 올바르게 평가할 수 있는 토대를 마련함으로써 기업의 경영성과라는 지식체계를 기업 내·외부의 전 조직이 공유할 수 있는 이론적 틀을 제시하였다.

그리고 이러한 이론 체계를 물리적으로 구현하기 위하여 경영성과 데이터웨어하우스를 구축하였고, 경영성과의 변화를 주도하는 중요 관리변수를 추출하기 위하여 유전알고리즘을 활용한 데이터마이닝체계를 구현하였다. 특히 경영성과 데이터베이스내의 데이터가 갖는 비정형적 특성을 고려하여 예측력 · 타당성 · 이해력을 동시에 고려한 설명기반 유전알고리즘체계를 제시하였다.

이와 같은 체계를 구축함으로써 기존과 같이 단발적인 방법으로 기업의 경영성과를 분석하는 것이 아니라, 경영성과의 변화를 연속적으로 추적하고 이에 영향을 미치는 여러 가지 변수를 실시간으로 분석함으로써, 기업 내 · 외부의 모든 이해관계자가 경영성과에 대한 정보 · 지식을 공유할 수 있는 체계를 구축할 수 있다고 판단된다. 그리고 구축된 경영성과에 대한 정보 · 지식의 발견 · 공유체계는 급격한 변화에 능동적으로 대처할 수 있는 바람직한 의사결정의 토대가 될 수 있을 것이라 생각된다.

6. 참고문헌

- [1] 김광용, "여러 가지 Data Mining 기법으로부터 도출된 지식에 관한 전문가의 신뢰도에 대한 실증적 연구", 한국지능정보시스템학회논문지, 제5권, 제1호, pp.125-143, 1999.
- [2] 김철중, "경영성과지표로서의 경제적부가가치의 유용성에 관한 연구", 재무관리논총, 제2권, 제1호, pp.101-126, 1995.
- [3] 김태성, "생산성관리를 위한 네트워크모델 구축에 관한 연구", 건국대학교 대학원 박사학위논문, p.50, 1997.
- [4] 김현수 · 이창호, "데이터웨어하우스 환경에서의 설명기반 데이터마이닝", 한국지능정보시스템학회논문지, 제5권, 제2호, pp.15-27, 1999.
- [5] 지천삼, "경제적부가가치(EVA)와 기업경영", STOCK, No.339, Nov., pp.3-34, 1996.
- [6] Bo K. Wong, Thomas A. Bodnovich and Yakup Selvi, "Neural Network Applications in Business : A Review and Analysis of the Literature(1988-95)", Decision Support System, 19, pp.301-320, 1997.
- [7] Guido J. Deboeck, Trading on the Edge(Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets), John Wiley & Sons, Inc., pp.244, 1994.
- [8] Paul A. Strassmann, Information Productivity, The Information Economics Press, pp.37, 1999.
- [9] Ray Tsaih, Yenshan Hsu and Charles C. Lai, "Forecasting S&P 500 stock index futures with a hybrid AI system", Decision Support Systems, 23, pp.166-174, 1998.
- [10] Shawn Tully, "The Real Key to Creating Wealth," Fortune, September 20, pp.38-50, 1993.
- [11] Usama Fayyad, "From Data Mining to Knowledge Discovery in Databases", AI Magazine, American Association for Artificial Intelligence, pp.37-54, 1996.