

데이터베이스 지식발견체계에 기반한 경영성과 정보시스템의 구축

조성훈¹ · 정민용² · 김종화²

¹삼성전자 경영혁신팀 / ²전국대학교 산업공학과

Modeling a Business Performance Information System with Knowledge Discovery in Databases

Seong-Hoon Cho¹ · Min-Yong Chung² · Jong-Hwa Kim²

We suggest a Business Performance Information System with Knowledge Discovery in Databases(KDD) as a key component of integrated information and knowledge management system. The proposed system measures business performance by considering both VA(Value-Added), which represents stakeholder's point of view and EVA(Economic Value-Added), which represents shareholder's point of view. In modeling of Business Performance Information System, we apply the following KDD processes : Data Warehouse for consistent management of a performance data, On-Line Analytic Processing(OLAP) for multidimensional analysis, Genetic Algorithms for exploring and finding dominant managing factors and Analytic Hierarchy Process(AHP) for applying expert's knowledge and experience. To demonstrate the performance of the system, we conducted a case study using financial data of Korean automobile industry over 16 years from 1981 to 1996, which is taken from database of KISFAS (Korea Investors Services Financial Analysis System).

1. 서론

급속하게 변화하는 경영환경에서의 생존이라는 측면에서, 기업의 경영관리영역인 재무(Finance) · 생산(Manufacturing) · 판매(Marketing) 등에서 발생하는 기업의 현황을 자동화된 정보 체계로 흡수 · 축적하여 조직 전체가 공유할 수 있는 정보 · 지식으로 변환하는 체계를 구축하는 것이 기업의 생존을 결정하는 매우 중요한 요소로 부각되고 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 역할을 기업 목표라는 초점에 집중시킬 수 있는 기반을 갖고 있어야 한다.

이에 본 연구는 변화하는 기업환경에 부응하기 위하여, 경영성과지표를 관리하는 의사결정자의 관점에서 정보기술의 효과적인 활용을 통하여 기업 전체의 조직이 공유할 수 있는 경영성과 정보시스템을 전개하고자 한다.

본 연구에서는 데이터베이스 지식발견체계(KDD : Knowledge Discovery in Databases)속에서 경영성과 정보시스템을 구체화한다 (Deogun, 1997; Fayyad, 1996) 제시된 경영성과 정보시스템은 기존의 단발적인 방법으로 기업 경영성과를 분석하는 것이 아니라, 경영성과의 변화를 연속적으로 추적하고 이에 영향을 미치는 여러 가지 변수를 실시간으로 분석할 수 있으며 기업 내 · 외부의 모든 이해관계자가 경영성과에 대한 정보 · 지식을 공유할 수 있는 물리적 토대를 제공할 수 있다. 그리고 구축된 경영성과에 대한 정보 · 지식의 발견 · 공유체계는 급격한 변화에 능동적으로 대처할 수 있는 바람직한 의사결정의 토대가 될 수 있을 것이라 판단된다.

2. 이해관계자 관점과 가치경영 관점을 결합한 경영성과

2.1 부가가치 중심의 경영성과

부가가치 구성항목에는 경상이익, 인건비, 금융비용, 임차

료, 조세공과 및 감가상각비가 포함된다. 이는 기업의 성과로서 창조가치 총액인 동시에 기업성과에 기여한 모든 이해집단에 대한 성과배분을 나타내기도 한다. 그러므로 부가가치가 증대되면 그만큼 기여집단에 대한 분배액도 커지며, 기업·산업·국가경제와 더불어 근로자의 생활수준도 보다 더 향상될 수 있다. 특히, 최근 지식이나 정보와 같이 기업이 가진 무형자산이 중시되는 환경에서는 자기자본수익률이나 투자수익률보다 부가가치를 대가로 한 이익률이 기업의 수익률을 더욱 적절하게 반영할 수 있다. 또한 조직이 고객의 요구를 얼마나 만족시키는가에 대한 효과성을 평가하는 경우에도 부가가치가 가장 좋은 대안으로 지목되고 있다. 결국 부가가치는 지식기업의 생산능력과 전통적인 재무지표의 단점을 보완하는 측정지표로 각광받고 있다 (Sveiby, 1999).

한국은행의 '기업경영분석'에서 사용하고 있는 가산법을 중심으로 한 부가가치는 다음의 식 (1)과 같다.

$$\begin{aligned} \text{부가가치} &= \text{경상이익} + \text{인건비} + \text{금융비용} + \text{임차료} \\ &+ \text{조세공과} + \text{감가상각비} \end{aligned} \quad (1)$$

2.2 경제적 부가가치 중심의 경영성과

경제적 부가가치는 미국의 컨설팅회사 Stern Stewart 사가 1980년대 후반에 도입한 개념으로서, 기업이 영업활동을 통해 창출한 순가치의 증가분인 세후순영업이익에서 투입된 자본에 대한 자본코스트를 차감한 금액이다. 과거에 중시되던 당기순이익을 기준으로 한 경영지표는 타인자본에 대한 비용만을 반영하고 있으므로 자기자본비용을 반영하지 못했다. 이에 반해 EVA는 기회비용이 반영된 주주의 기대수익을 자본비용에 포함시킴으로써 주주의 기회비용을 명확히 할 수 있다. 또한 EVA는 기업이 영업활동을 통해 달성한 이익에서 이를 위해 투입된 자본에 대한 정당한 자본비용을 차감한 후에 성취된 이익의 양을 측정하는 것이다. 이러한 의미에서 EVA는 기업의 근본적인 생산활동만을 그 대상으로 삼고 있다고 할 수 있다. EVA의 산출과정은 정리하면 다음의 식 (2)와 같다 (山一經濟研究所, 1995; Tully, 1993).

$$\begin{aligned} \text{EVA} &= \text{세후순영업이익(Net Operating Profits After Taxes)} \\ &- \text{자본코스트} \\ &= (\text{영업용투자자본} \times \text{투자자본수익율}) \\ &- (\text{영업용투자자본} \times \text{자본비용}) \\ &= \text{영업용투자자본} \times (\text{투자자본수익율} - \text{자본비용}) \end{aligned} \quad (2)$$

2.3 부가가치와 경제적 부가가치를 결합한 경영성과

현행의 회계이익으로 경영성과를 평가하는 과정에서는 자기자본조달사용에 대한 대가로 지불되어야 할 자기자본비용을 비용으로 인식하지 않는다. 따라서 단순히 회계상의 경상

이익을 기준으로 경영성과를 평가하는 것은 자본제공자인 주주의 입장에서 판단한다면 자본조달에 따른 자본비용을 충분히 감안하지 못한다고 인식할 수 있으며, 사전에 예상된 자기자본비용 보상에 대한 주주의 중요 관심사를 외면하게 된다.

국내의 연구 결과, 1987년부터 1992년까지 EPS(Earning Per Share)가 모두 양의 값을 갖는 표본상장기업 중 약 70~80%에 이르는 기업들의 「경제적 부가가치/주(Share)」가 음(-)의 값을 갖는 것으로 나타났다. 미국의 경우에도 양의 회계이익을 기록하면서도 음의 경제적 부가가치를 나타내는 기업이 약 21%나 됨으로써 상당수의 기업들이 영업활동을 통하여 얻는 수익성으로 자기자본조달비용을 충분히 보상하지 못하고 있다고 판단할 수 있다 (김철중, 1995).

이와 같이 기업이 산출한 이익 개념으로서 회계상의 당기순이익보다는 영업관련 현금흐름에 입각한 경제적 부가가치가 중시되는 최근의 경향을 고려하여, 본 연구에서는 기업의 경영성과를 평가하는 틀로서 이해관계자 입장에서의 1인당부가가치(부가가치/종업원수)와 자본 제공자 입장에서의 투자자본대비 경제적 부가가치(EVA/평균투자자본)를 동시에 사용하고자 한다. 이와 같은 논의를 종합하여, 본 연구에서 제시하는 경영성과분석을 위한 구조를 제시하면 다음의 <그림 1>과 같다.

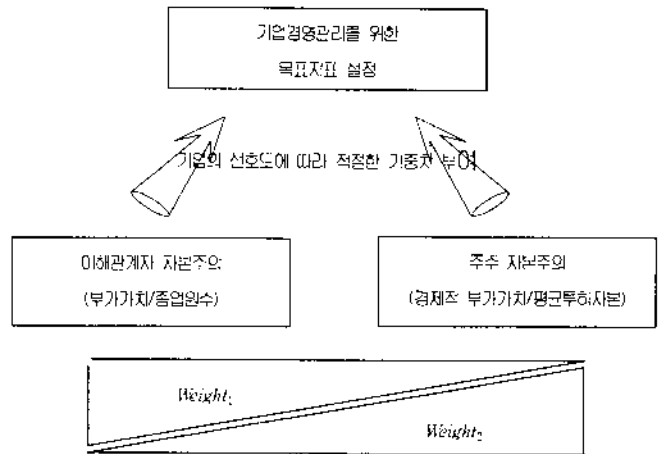


그림 1. 경영성과분석을 위한 목표지표 설정.

2.4 비재무적요인을 고려한 경영성과

재무제표를 바탕으로 한 정량적 분석만으로는 완전한 경영성과분석이 어렵다고 할 수 있다. 기업의 경영성과를 평가하기 위해서는 부가가치, 현금흐름, 경제적 부가가치 등의 재무관련 정량적 지표와 기업의 가치창출원동력 등과 같은 비재무적 요인을 동시에 고찰해야 하기 때문이다. 가치창출 원동력에는 불량률, 재고관리수준, 소비자만족도, 기업문화 등과 같은 여러 가지 비재무적 요소가 포함될 수 있다.

특히 Kaplan과 Norton이 제시한 균형잡힌 성과측정기표(Balanced Scorecard: BSC)에서는 재무관점의 평가지표로서 단기적인 성과를 측정하는 동시에, 상호 보완적인 측면에서 고객

(Customers)관점 · 내부프로세스(Internal Business Process)관점 · 학습 및 성장(Learning and Growth)관점 등을 통한 성과측정 지표를 제시하였다 (Kaplan, 1996).

또한 Strassmann은 기존에 사용되어 왔던 제품원가 · 종업원 수 · 매출액 · 수익 · 이익 · 자산 규모 등의 경영성과지표가 현실적인 경제현상을 반영하지 못하며, 정보기반의 경제가 자본기반의 경제를 추월하는 현실을 고려할 때, 경영성과나 기업가치를 측정하기 위해서는 정보기반의 생산성을 측정 · 활용하여야 한다고 주장한 바 있다 (Strassmann, 1999).

본 연구에서는 자료구성의 한계상 불량률, 재고관리수준, 소비자만족도, 기업문화, 정보기술수준 등과 같은 비재무적 요소는 고려하지 않는다. 그러나 포괄적인 기업경영관리 체계의 필요성과 더불어, 재무적 정보와 함께 비재무적 정보도 그 중요성을 더해가는 현실을 감안한다면 이러한 한계점은 반드시 보완되어야 할 것이라 판단된다.

3. KDD에 기반한 경영성과 정보시스템의 이론모델 구축

1인당 부가가치와 투자자본대비 경제적 부가가치를 결합한 경영성과평가의 이론구조를 바탕으로, 본 연구에서 구현하는 데이터베이스 지식발견체계에 기반한 경영성과 정보시스템의 활용체계를 요약하면 다음의 <그림 2>와 같다.

경영성과데이터의 지속적 관리차원에서 데이터웨어하우스(Data Warehouse)와 해당 데이터의 다차원적 분석차원에서 OLAP(On-Line Analytic Precess)를 도입한다. 또한 경영성과의 변화를 주도하는 중요 관리변수를 추출하기 위하여 유전알고리즘(Genetic Algorithms)을 활용한 데이터마이닝(Data Mining)체계를 구현하고, 기계학습의 귀납적 오류를 보완하고 또한 전문가의 사전지식을 활용한다는 측면에서 계층적 의사결정 기법인 AHP(Analytic Hierarchy Process)를 도입하여 중요 관리변수를 추출할 수 있는 체계도 도입한다.

이와 같은 시스템을 통하여 조직 내외의 이해관계자는 경영

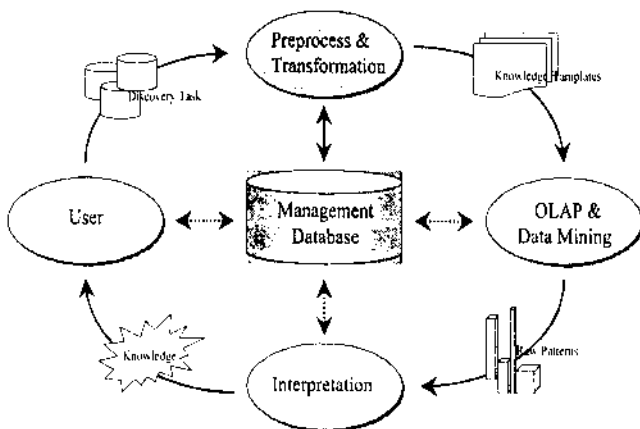


그림 2. 데이터베이스 지식발견체계에 기반한 경영성과 정보시스템의 활용 체계.

성과에 대한 Discovery Task(지식발견의 구체적 대상)를 지정하고 경영성과 데이터베이스에서 Knowledge Templates(지식발견을 목적으로 사전처리된 목표 Data)을 생성한다. 그리고 클라이언트 내에 포함된 OLAP/Data Mining 기능을 이용하여 경영성과에 대한 새로운 경향, 패턴 그리고 규칙 등을 생성할 수 있으며, 차트나 산점도 등의 시각화(Visualization) 도구를 활용한 해석을 통하여 구체적인 경영성과 지식을 도출할 수 있다.

3.1 경영성과 데이터웨어하우스의 구성

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

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」, 「기업계반」 테이블을 구성하고 이를 기업의 경영성과분석을 위한 원시데이터로서 사용한다.

이러한 5개의 운영데이터(Operational Data) 테이블을 기반으로 「성장성」, 「규모」, 「수익성(수익/자본)」, 「수익성(수익/비용)」, 「수익성(비용구조)」, 「안정성」, 「활동성」, 「경제적 부가가치」, 「생산성」의 9개 정보데이터(Information Data) 테이블을 생성 · 유지한다. 관리자 이외의 분석 목적으로 시스템에 접근하는 사용자들은 9개의 정보데이터 테이블에 접근하여 기업의 경영성과를 분석한다. 구축된 데이터베이스의 개체-관계도(Entity-Relationship Diagram)를 작성하면 다음의 <그림 3>과 같다.

<그림 3>의 개체-관계도는 각 사업부에 대하여 관리되어야 할 데이터들을 그 특성에 맞게 분류하여 Data Modeling한 것이며, 범용성을 고려하여 각 데이터 테이블이 반드시 필수적인 것은 아니도록 설계하였다. 예를 들어, 사업부는 한 연도 이상의 대차대조표들을 가지거나 갖지 않을 수도 있지만, 대차대조표가 존재하는 경우 그 사업부는 반드시 존재해야 한다. 관리자는 서버를 통하여 원시데이터베이스로서 5개 운영데이터 테이블에 경영성과 자료를 입력하며, 클라이언트를 통하여 사용자의 요청이 발생하면 운영데이터 테이블을 기반으로 9개의 정보데이터 테이블이 실시간으로 생성된다. 실제로 클라이언트에서는 9개의 정보데이터 테이블에만 접근이 허용된다.

3.2 경영성과 큐브를 통한 OLAP

본 연구에서는 전술한 관계형 데이터베이스 구조의 제약극복하기 위하여 데이터 큐브를 사용한다. 데이터 큐브를 이

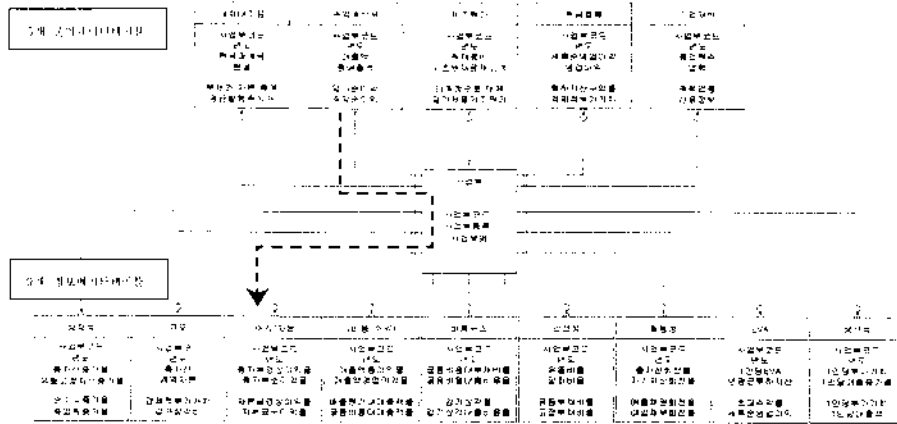


그림 3. 경영성과 데이터베이스의 개체-관계도.

용하면 분석하는 데이터의 양에 크게 영향을 받지 않고 동일한 성능으로 데이터를 분석할 수 있으며, 다차원적 데이터 분석 과정을 지원하는 과정에서 단순 수치의 나열이외에도 추가적인 연산과정을 지원할 수 있다.

경영성과 데이터 큐브는 「시간, 「조직단위, 「재무항목, 「재무비율(값)의 애트리뷰트들을 가지며, 여기에서 차원 애트리뷰트들은 「시간, 「조직단위, 「재무항목」이며, 이 애트리뷰트 값들의 조합이 측정 애트리뷰트 「재무비율(값)의 한 값을 결정한다. 따라서 본 연구에서 사용되는 3-차원 데이터 큐브는 차원 애트리뷰트로서 *Date*, *Department*, *Item*을 좌표축으로 가지며 측정 애트리뷰트로서 *Ratio(Vaule)*의 값에 대응되는 좌표의 원소값을 가지는 다차원 배열로 나타남을 알 수 있다. 이를 도시하면 다음의 <그림 4>와 같다.

3.3 유전알고리즘을 활용한 경영성과의 중요 관리변수 추출

과거 자료 내에 존재하는 변수들 간의 관련성을 분석하는 도구로써 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 데이터베이스 내의 재무관련 비율은 여러 요소가 중복적으로 비율의 분모와 분자로 사용되므로, 다중공선성의 문제가 자주 발생하기 때문에 통계적 분석기법을 사용하기에는 여러 가지 제약이 존재한다.

이러한 한계를 감안하여 본 연구에서는 기계학습의 한 종류인 유전알고리즘을 통하여 경영성과의 중요 관리변수를 추출한다. 다양한 기계학습론 중에서 유전알고리즘을 활용하는 이유는 통계적 분석기법과는 달리 선형성·정규성의 가정이 필요하지 않으며, 해공간이 넓고 자료의 노이즈가 심한 경우에도 전역해를 빠른 시간에 탐색할 수 있기 때문이다. 그러나 일반적인 유전알고리즘은 무작위로 발생된 난수값을 기반으로 평가·선택·교차·돌연변이 등의 연산자를 거치면서 추출된 귀납적 결과를 받아들여야 하므로, 반/비 정형화된 문제의 경우에는 의사결정자가 그 결과를 전적으로 받아들여기가 쉽지 않다(김현수, 1999; Deboeck, 1994; Hart, 1992).

본 연구에서는 이러한 측면을 고려하여 유전알고리즘을 선형결합형태로 설계함으로써, 복잡한 비선형구조에 비하여 예측도는 떨어질 수 있으나 입력변수와 결과변수 간의 선형가중치를 변수의 중요도로 그대로 사용하여 도출된 지식의 이해력(결과의 의미)을 증진시킨다. 또한 유전알고리즘의 초기 유전자값을 일양분포 난수로 발생시키지 않고, 분석자료입력시 산출되는 입력변수와 결과변수의 상관계수(Correlation Coefficient)와 결정계수(Determination Coefficient)를 활용한 정규분포 난수로 발생시킴으로써 추출된 지식의 타당성(결과의 수용도)을 높인다.

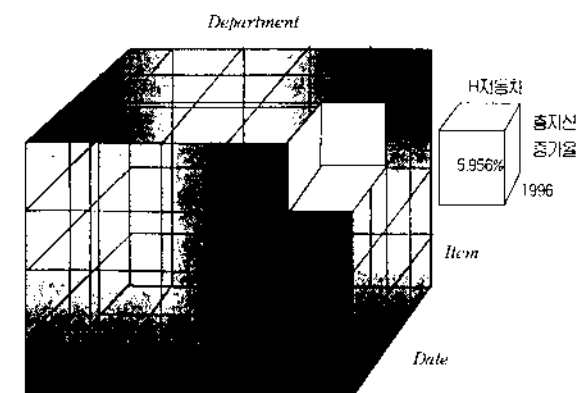


그림 4. 경영성과 데이터 큐브.

3.3.1 선형결합구조의 설계

본 연구에서 제시되는 중요변수추출을 위한 유전알고리즘은 선형결합구조를 취하는데, 이는 선형결합시 사용되는 가중치를 변수의 중요도로 활용하기 위해서이며, 이는 <그림 5>와 같다.

이와 같은 선형결합구조에서는 선형계수를 변수의 가중치로 사용함으로써, 비선형모형이 갖는 Black-Box 특성을 지양하

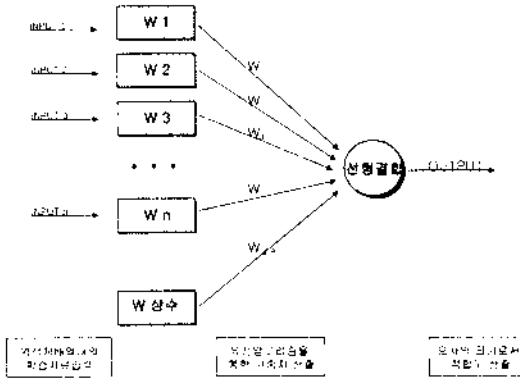


그림 5. 유전알고리즘 내의 선형결합구조.

고 그 결과를 쉽게 분석할 수 있으며, 중요 관리변수의 추출 결과에 대한 이해력을 높일 수 있다.

3.3.2 정규분포를 이루는 난수구조의 설계

상관분석은 두 변수 간의 선형관계의 방향과 정도를 나타내는 통계적 분석기법으로, 상관계수를 이용하면 입력변수로서의 목표변수와 결과변수로서의 관리변수 간의 관계를 쉽게 이해할 수 있다. 또한 결정계수는 관리변수가 목표변수의 변동을 설명하는 비율이므로, 상대적인 중요성의 판단기준으로 사용할 수 있다.

따라서 본 연구에서 제시하는 유전알고리즘에서는 상관계수로써 초기 유전값의 방향을 결정하고, 결정계수로써 유전값의 크기를 결정한다. 결정계수값을 정규분포의 평균값으로 활용함으로써, 유전알고리즘의 무작위 발생 특성은 그대로 살리면서 결정계수 근처의 값이 많이 발생하도록 연산자를 설계한다. 이로부터 통계적 가정에는 강건(Robust)하면서도 통계적 특성이 결합된 유전알고리즘을 수행할 수 있다.

정규분포 난수 발생을 위해서는 반드시 수치해석법(Numerical Method)을 사용하여야 하는데, 이는 정규분포의 누적분포나 역함수를 정확하게 나타내는 함수가 없기 때문이다. 수치해석법으로써 Beasley · Springer가 제시한 방법을 사용한다(Beasley, 1977).

이 방법에서는 일양난수 p 에 대하여 $ABS(p-0.5)$ 가 0.42보다 작은 경우에는 식 (3)을 사용하고, 반대로 $ABS(p-0.5)$ 가 0.42보다 큰 경우에는 식 (4)를 사용하여 정규분포 난수를 발생시킨다.

$$Random(p) = (p-0.5) \times \frac{A(p)}{B(p)} \quad (3)$$

$$Random(r) = \pm \frac{C(r)}{D(r)}, r = \sqrt{\ln(0.5) - ABS(p-0.5)} \quad (4)$$

여기서 $A(p)$, $B(p)$, $C(r)$, $D(r)$ 은 각각 p 의 3차 다항식, p 의 4차 다항식, r 의 3차 다항식, r 의 2차 다항식이다.

3.3.3 학습구조 설계

유전알고리즘을 활용한 중요 관리변수의 학습구조를 도시

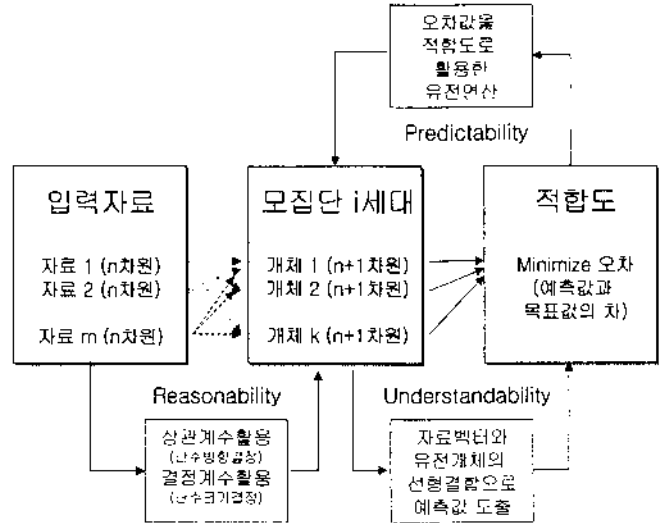


그림 6. 중요 관리변수 추출을 위한 유전알고리즘 학습구조.

하면 다음의 <그림 6>과 같다.

<그림 6>의 학습구조에서는 비선형모형으로써 예측력의 증대를 추구하지 않고 선형구조를 채택함으로써 추출된 중요 관리변수의 이해력을 도모하였으며, 정규분포를 활용한 초기 유전값 설계과정을 통하여 지나친 탐색공간의 확대를 방지하고 또한 상관계수와 결정계수를 초기 유전값의 정규난수 Seed로 사용함으로써 기계학습의 결과를 합리적으로 받아들일 수 있는 타당성 측면을 고려하였다. 이와 같은 설명기반 유전알고리즘 체계는 예측력(Predictability)이라는 요인 이외에도 이해력(Understandability) · 타당성(Reasonability)이라는 특성을 동시에 추구한다고 할 수 있다.

3.4 계층화 의사결정을 활용한 경영성과의 중요 관리변수 추출

전술한 바와 같이 아무런 사전지식 없이 기계학습 등의 데이터마이닝기법을 통하여 생성된 결과는 가늠적 오류를 범할 가능성이 존재한다. 따라서 대량의 자료를 분석하는 과정에서, 효과적인 지식 발견을 추구하기 위해서는 인간이 갖고 있는 지식을 적절하게 결합하는 것이 매우 중요한 문제가 될 수 있다. 따라서, 정성적인 요인을 정량화하여 기존의 정량적인 모형과 결합할 수 있는 체계가 요망되며 본 연구에서는 그 도구로써 AHP를 사용하고자 한다.

본 연구에서는 2장에서 서술한 바와 같이 기업의 경영성과 분석 목표로서 1인당 부가가치와 투자자본대비 경제적 부가가치라는 두 가지 요인을 결합하여 사용한다. 따라서 의사결정계층에서 1인당 부가가치와 투자자본대비 경제적 부가가치가 결합된 목표변수가 계층의 첫 번째 수준이 된다. 그리고 9개의 정보데이터 테이블이 두 번째 수준이 되며, 정보데이터 테이블 내에 있는 모든 변수들이 세 번째 수준을 구성한다. 이와 같은 개념을 도시하면 다음의 <그림 7>과 같다.

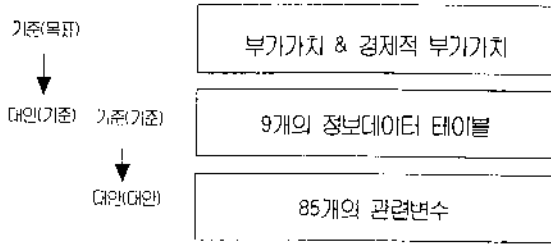


그림 7. 수정된 중요변수 추출의 계층.

3.5 경영성과 정보시스템을 위한 KDD 프로세스

본 연구에서 구현하는 데이터베이스 지식발견체계의 목적과 구성요소를 정의하면 다음의 <표 1>과 같다.

표 1. 경영성과 정보시스템의 목적과 구성요소

용어	정의
목적	경영성과의 변화를 연속적으로 추적하고 이에 영향을 미치는 85개 관리변수를 실시간으로 분석함으로써, 기업 내·외부의 모든 이해관계자가 경영성과에 대한 정보·지식을 공유할 수 있는 체계를 구축
데이터웨어하우스	기업의 제조·판매를 통하여 발생한 재무관련 자료를 내/외부 자료원으로부터 자동적으로 추출하여 경영성과분석에 대한 의사결정 환경에 맞게 최적화하여 통합화하는 주제 지향적인 읽기 전용 데이터베이스
데이터웨어하우징	운영데이터(Operational Data)에서 정보데이터(Information Data)를 추출, 변환하여 데이터웨어하우스에 적재하는 총괄적인 프로세스
OLAP	조직단위, 시간단위 분석에 필요한 데이터를 통합, 추출함으로써 계열적인 분석을 다차원적으로 실시간에 처리하는 프로세스
데이터 마이닝	의사결정의 효율성을 높이고 재무분석의 결과를 현장관리와 연계하기 위하여 경영성과와 관련된 85개의 관리변수 중 상대적으로 중요한 관리변수를 추출하는 프로세스

이와 같은 논의를 통하여 종합된 본 연구의 KDD Process는 다음의 <그림 8>과 같이 요약된다.

4. 경영성과 정보시스템의 구축 및 활용

본 연구에서 구현한 경영성과 정보시스템은 관계형 데이터베이스를 기반으로 한 클라이언트-서버 구조에 바탕을 두고 있다. 서버의 운영시스템으로 윈도우 NT 4.0을 사용하고 데이터웨어하우스는 MS-SQL 7.0을 사용해 원시 데이터베이스와 같은 시스템상에서 구현하였다. Mining서버는 윈도우 NT상에서

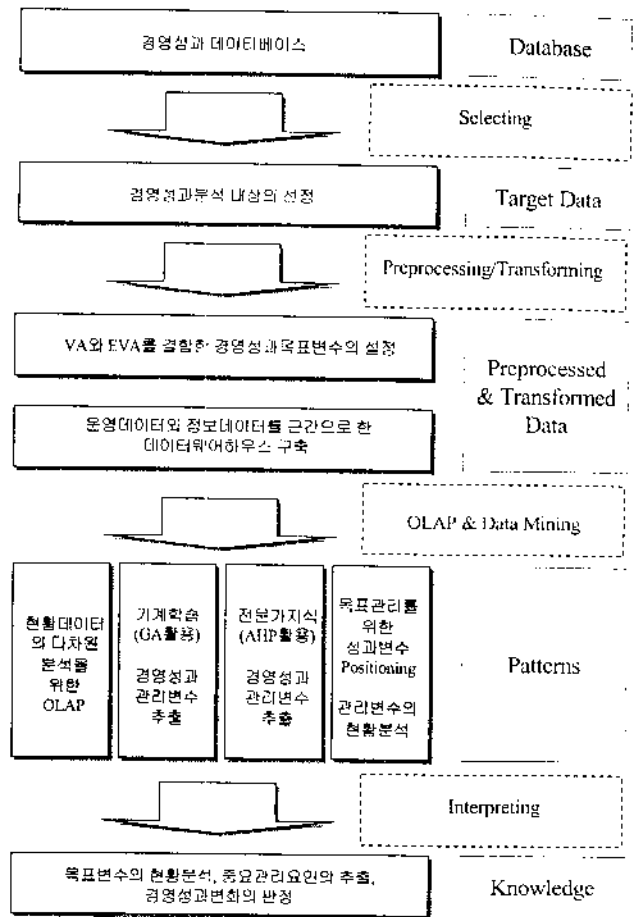


그림 8. 경영성과 정보시스템의 KDD 프로세스.

구현하며 MS-Visual Basic 6.0을 통하여 구축되었으며, Mining 클라이언트는 윈도우 98상에서 구현하며 역시 MS-Visual Basic 6.0을 통하여 구현되었다. 서버 관리자는 MS-SQL 7.0의 데이터 변환서비스(Data Transformation Service)를 이용하여 원시데이터를 운영데이터 테이블에 전송할 수 있다.

4.1 경영성과 정보시스템의 구성 및 기능

데이터베이스 지식발견체계에 기반한 경영성과 정보시스템의 구성 및 기능은 다음의 세가지 측면에서 전개될 수 있다.

- 데이터 접근범위
- 목표/관리변수 VIEWER
- 중요변수추출

데이터 접근범위는 클라이언트를 통하여 서버에 접근하는 User에게 어느 정도의 권한을 부여할 것인가를 결정하는 것이다. 서버 관리자 수준에서는 사용자 수준과는 달리 원시데이터에 대한 직접적인 조작이 가능하다.

목표/관리변수 VIEWER에는 목표변수 POSITIONING 기능, 목표변수 시계열 기능, 관리변수 시계열 기능 등이 포함되는데, 목표변수 POSITIONING은 여러 조직 간의 경영성과를 판

별하는 기능을 수행하며, 목표변수 시계열은 개별 조직 또는 여러 조직 간의 경영성과를 시간축에 따라 표시함으로써 해당 조직의 경영성과 변동여부를 파악할 수 있다. 또한 관리변수 시계열은 목표변수의 구체적 변화를 반영하는 85개 재무비율의 변화를 파악하는 체계로서 개별 조직 또는 여러 조직 간의 관리변수 추세를 비교·분석할 수 있다.

중요변수 추출에는 목표/관리변수 OLAP, 유전알고리즘을 이용한 데이터마이닝, 계층화 의사결정기법을 이용한 가중치 산출 등이 포함된다. 목표/관리변수 OLAP는 실시간의 경영성과 시계열분석과 함께, 조직단위 간의 횡단면적 분석도 가능하며, 시계열과 횡단면적 분석을 동시에 다차원적으로 실시할 수도 있다. 유전알고리즘을 이용한 데이터마이닝은 많게는 85개 차원에 대하여 관리변수의 가중치를 동시에 추정함으로써 목표변수 변화에 민감하게 반응하는 관리변수를 기계학습으로써 추출할 수 있으며, 이로부터 조직의 관리역력을 중요 관리변수에 선택·집중함으로써 그 효과를 극대화할 수 있다. 계층화 의사결정기법을 이용한 가중치 산출 기능은 귀납적 오류를 가질 수 있는 유전알고리즘 체계를 보완하며 전문가의 사전 지식을 활용하여 관리변수의 중요도를 추출한다. 이와 같은 기능 전개를 종합하여 도시하면 다음의 <그림 9>와 같다.

이와 같은 기능 전개를 통하여 시스템 사용자는 기업 조직의 경영성과를 체계적으로 분석하고 그 경영성과를 구체적으로 관찰할 수 있으며, 외부 접근을 통하여 기업의 경영성과 결과를 전 조직원이 공유할 수 있다.

4.2 경영성과 정보시스템의 활용

본 연구에서 제시하는 경영성과 정보시스템의 목적은 기업에 포함된 여러 조직단위간의 경영성과의 정보 및 지식을 최고 의사결정자는 물론 자기자본 소유주·타인자본 소유주·종업원·정부 등의 기업 내·외부의 이해관계자 전체가 공유함으로써, 기업을 올바르게 유지시킬 수 있는 효과적인 경영

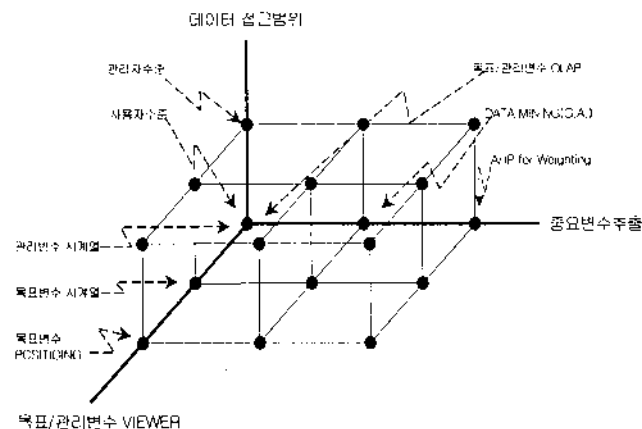


그림 9. 데이터베이스 지식발견체계에 기반한 경영성과 정보시스템의 기능 전개.

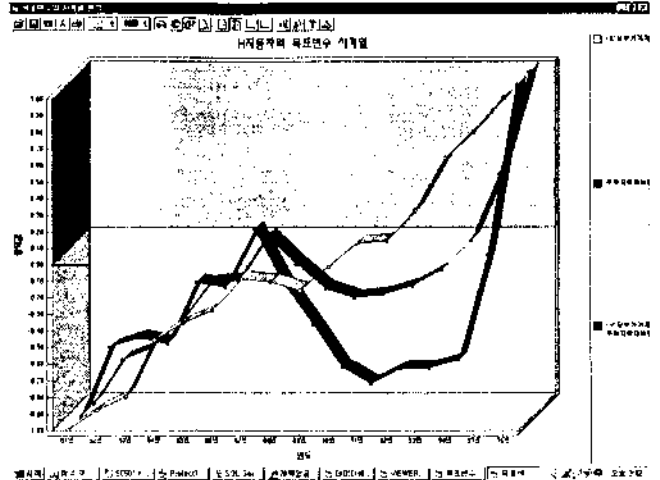


그림 10. H자동차의 목표변수 시계열 시각화.

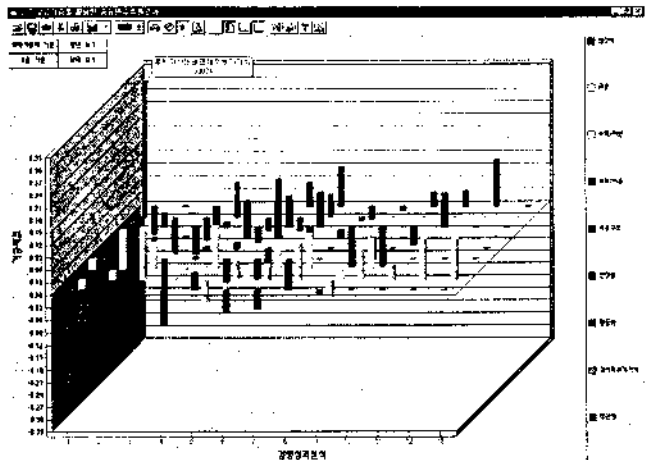


그림 11. 유전알고리즘을 활용한 중요 관리변수 시각화.

의사결정을 유도하는 것이다.

그러나 공개된 자료의 한계상, 동일 기업 내의 여러 조직단위(사업부 등)의 경영성과를 대상으로 적용 사례를 전개하지 못하고, 기업 간의 자료로서 국내 자동차업종을 그 대상으로 선정하였다. 기업 경영성과분석의 기준이 되는 5개 운영데이터 테이블의 자료는 (주)한국신용평가의 재무제표 데이터베이스인 KISFAS(Korea Investors Services Financial Analysis System)를 활용하였으며, 국내의 자동차산업을 대상으로 1981-1996년까지의 자료를 이용하였다.

경영성과 정보시스템의 여러 기능 중 목표변수의 시계열 결과와 설명기반 유전알고리즘을 이용한 마이닝 결과를 예시하면 <그림 10>, <그림 11>과 같다.

5. 결론 및 향후 연구

본 연구에서는 변화하는 기업환경에 부응하기 위하여, 경영성과 지표를 관리하는 의사결정자의 관점에서, 정보기술의 효과

적인 활용을 통하여 기업 전체의 조직이 공유할 수 있는 경영성과분석 체계를 전개하였다. 이를 위하여 기업의 경영성과분석의 틀을 이해관계자 관점에서의 부가가치와 가치경영 관점에서의 경제적 부가가치로서 제시하였으며, 제시된 경영성과분석 체계를 통하여 기업 내부의 이해관계자는 물론 기업 외부의 이해관계자 역시 기업의 경영성과를 올바르게 평가할 수 있는 접근 체계를 제시하였다.

그리고 이러한 이론 체계를 데이터베이스에 기반한 지식발견체계 속에서 물리적으로 구현하였다. 이를 위하여 경영성과데이터의 지속적 관리차원에서 데이터웨어하우스와 해당데이터의 다차원적 분석차원에서 OLAP를 도입하였다. 또한 경영성과의 변화를 주도하는 중요 관리변수를 추출하기 위하여 유전알고리즘을 활용한 데이터마이닝체계를 구현하였고, 기계학습의 귀납적 오류를 보완하고 또한 전문가의 사전지식을 활용한다는 측면에서 계층적 의사결정 기법인 AHP를 도입하여 중요 관리변수를 추출할 수 있는 체계도 도입하였다.

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.

향후 연구과제로는 신용등급과 같이 기업의 경영성과 정도를 절대적으로 판단할 수 있도록 지표를 확인하는 작업과 이를 위하여 좀더 많은 자료를 데이터베이스로 축적하는 작업으로 요약할 수 있다. 특히 비재무적 요소와 함께, 상부의 기업경

영성과 하부의 물적변환과정과의 연계성을 파악하여 이를 구체적으로 데이터화하는 것이 중요한 향후 연구과제로 판단된다.

참고문헌

김철중 (1995), 경영성과지표로서의 경제적 부가가치의 유용성에 관한 연구, *재무관리논총*, 2(1), 101-126.
 김현수, 이창호 (1999), 데이터웨어하우스 환경에서의 설명기반 데이터마이닝, *한국지능정보시스템학회논문지* 5(2), 15-27.
 山一經濟研究所 (1995), 企業財務の近代化とEVA, *證券月報*, 5, 4-39.
 Beasley, J. D. and Springer, S. G. (1977), The Percentage Points of the Normal Distribution, *Applied Statistics*, 26, 118-121.
 Deboeck, G. J. (1994) *Trading on the Edge* (Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets), John Wiley & Sons, Inc., 1994.
 Deogun, J. S., Raghavan, V. V., Sarkar, A. and Sever, H. (1997), Data Mining : Trends in Research and Development, *Rough Sets and Data Mining*, Kluwer Academic Publishers, 9-46.
 Fayyad, U. (1996), From Data Mining to Knowledge Discovery in Databases, *AI Magazine*, American Association for Artificial Intelligence, Fall, 37-54.
 Hart, A. (1992.), *Knowledge Acquisition for Expert Systems* (2nd Ed.), New York : McGraw-Hill
 Inmon, W. H. (1996), *Building the Data Warehouse* (2nd Ed.), John Wiley & Sons, Inc.
 Kaplan, R. S. and Norton, D. P. (1996), Using the Balanced Scorecard as a Strategic Management System, *Harvard Business Review*, January-February, 75-85.
 Strassmann, P. A. (1999), Information Productivity(Assessing the Information Management Costs of U. S. Industrial Corporations), The Information Economics Press.
 Sveiby, K. E., 정선구·김용구 옮김 (1999), *지식경영 성공을 위한 지식자산의 측정과 관리*, 미래경영개발연구원.
 Tully, S. (1993), The Real Key to Creating Wealth, *Fortune*, 20 (9), 38-50.



조성훈
 건국대학교 산업공학과 학사
 건국대학교 산업공학과 석사
 건국대학교 산업공학과 박사
 현재: 삼성전자(주) 경영혁신팀
 관심분야: e-Business, SCM, Data-Mining



김종화
 서울대학교 산업공학과 학사
 서울대학교 산업공학과 석사
 University of Michigan Industrial and Operation Engineering 박사
 현재: 건국대학교 산업공학과 조교수
 관심분야: Material Handling, 물류, 시뮬레이션



정민용
 건국대학교 축산학 농학학사
 건국대학교 공업경영과 공학석사
 현재: 건국대학교 산업공학과 교수
 관심분야: 생산성, 연구방법론