

## LCD 산업에서의 품질마이닝 시스템 A Quality Data-Mining System in LCD Industry

이 현우<sup>1)</sup>, 남 호수<sup>2)</sup>, 최 병욱<sup>3)</sup>

<sup>1)</sup>(주)한국신뢰성기술서비스, 서울시 송파구 오금동 33-10 마스텀빌딩

<sup>2)</sup>동서대학교 시스템경영공학과, 부산광역시 사상구 주례2동 산 69-1

<sup>3)</sup>StatSoft Korea, 서울시 송파구 오금동 33-10 마스텀빌딩

### Abstract

본 논문에서는 LCD 공정의 효율적인 관리를 위한 방법으로서 품질마이닝 시스템의 설계/개발 그리고 운영방법론을 논하고자 한다.

주요내용으로는 주요공정의 탐색, 설비유의차분석, 공정최적화 및 recipe 최적화, 수율 및 주요특성의 추정/예측 등을 들 수 있다. 이를 위하여 다양한 데이터마이닝 도구와 통계적 모형의 적절한 활용 방법을 논하고자 한다. 또한, 실제현장 중심의 개발사례를 통하여 품질마이닝 시스템의 유용성을 기술하였다.

*Keywords* : LCD, 데이터마이닝, 품질마이닝, SPC, 마이닝시스템

### 1. 서론

21세기 들어 컴퓨터의 발달은 정보체계의 발전과 더불어 일상생활에 너무나 많은 변화가 일어나고 있는 게 사실이다. 우리나라 어떤 지역에 어떤 사건이 발생하였는가를 실시간으로 알아볼 수 있을 뿐만 아니라 전 세계 거의 모든 지역에서 일어나는 중요한 사건들도 거의 실시간으로 알아볼 수 있다. 예전에는 중요한 자료를 찾기 위해서는 대형 백과사전을 찾거나 큰 도서관에 가서 혹은 외국 도서관에 의뢰를 해서 정보를 찾아야만 했다. 그러나 지금은 그러한 것들을 인터넷 이라는 도구를 이용해서 훨씬 많은 정보를 자신의 책상에 앉아 아주 짧은 시간에 조사가 가능해졌다.

산업 전반에도 많은 변화가 일어났다. 집안에서

시장을 볼 수 있고, TV를 보면서 전화로 혹은 인터넷으로 옷을 구매할 수 있다. 회사에서는 이런 고객들의 성향을 파악해서 고객이 원하는 제품을 추천하고, 고객의 구매 패턴을 분석해서 마케팅에 활용한다. 이러한 활동을 하기 위해서는 수많은 고객들에 대한 정보와 그 고객들의 행동 하나하나를 데이터화 하고 분석하여야 한다. 그러나 이러한 분석은 불과 10년 전만 하더라도 그리 쉬운 일은 아니었다. 결국 컴퓨터의 발달이 이런 일을 가능하게 만들어 준 것이다.

이러한 활동은 제조업에서도 마찬가지로 일어나고 있다. 전 공정상에 상황을 자동으로 검사하고 이들 데이터를 DB화해서 실시간으로 전 공정의 상태를 파악할 수 있게 되었다. 또한 제품의 고장이나 결점의 상태를 파악하고 이들의 패턴을 감지함으로써 무결점의 완벽한 제품을 만들기 위한 노력들을 많은 시간을 들여 하고 있다. 고장이나 각종 결점에 대한 데이터를 통계분석하고 분석결과를 이용해서 결점의 원인을 해결하고, 다시 이러한 결점이 일어나지 않도록 예방보전을 해주는 과정을 벌이고 있다. 이런 방대한 데이터를 통계처리 할 수 있는 기법이 데이터 마이닝이다.

데이터마이닝의 이론이 최근 들어 새롭게 나온 것은 아니다. 이미 1960년 초부터 마이닝의 이론이 등장했으나 대량데이터의 정의와 이를 처리할 수 있는 컴퓨터의 용량 문제로 이를 제대로 활용할 수 없었으나, 최근 급속도로 발전하고 있는 IT의 발전이 이를 가능하게 하였다. 또한 약 10여년 전부터 급성장한 CRM은 데이터마이닝의 확산에 크게 기여하였다. 반면에 제조 공정에서의 마이닝의 적용은 CRM에서의 마이닝 보다 현실적으로 많은 어려움 때문에 그다지 활성화가 되어있지 않은 게

사실이다. 그러나 이러한 어려움에도 불구하고 무결점의 최상의 제품을 만들기 위해서는 반드시 필요한 과정이 데이터 마이닝이며 이를 효율적으로 활용하기 위해서는 데이터 마이닝 시스템의 구축이 필수불가결하다고 할 수 있을 것이다.

본 논문에서는 LCD분야에서 공정을 효율적으로 운영하기 위한 데이터마이닝 시스템에 대한 현황과 방법론을 설명하고 실제 구축된 사례를 통해서 구축내용과 활용사례를 설명하기로 한다.

## 2. LCD 공정에서의 데이터 마이닝 적용 현황

현재 LCD공정에서는 대부분 SPC시스템을 구축하여 고장이나 결점 현황들을 파악하고 조치하는 수준으로 이루어져 있으나 마이닝의 적용 사례는 그리 많지 않다. 현재까지 파악된 마이닝 적용의 몇 가지 문제점에 대해서 논의하기로 한다.

### 2.1 결측치

데이터 마이닝을 하려고 하는 가장 큰 목적은 결점의 원인이나 수율의 저하를 일으키는 원인을 보다 효율적으로 그리고 보다 빨리 찾아내서 빠른 대응을 통한 무결점 고 수율의 공정을 만들어 나가기 위함이다. 그러기 위해서는 하나의 제품이 만들어지는 전체 공정에서 공정 상태를 체크하고 이를 데이터화 한 후 이를 분석해서 원인을 찾아야 한다. 그러나 전체 공정에서 공정 상태를 측정하는 일은 그리 쉬운 일은 아니다. 측정을 한다는 것은 공정 중에 자동적으로 수많은 측정요소들을 측정할 수 있는 것이 아니기 때문에 이를 위해서는 공정 중에 측정을 위해 라인 밖에서 측정하여야 할 필요성도 있고, 경우에 따라서는 제품에 나쁜 영향을 미치면서 까지 측정을 실시하여야 한다. 따라서 하나의 제품이 만들어지는 과정 처음부터 마지막 공정 전체에서 측정이 이루어진다는 것은 거의 불가능에 가깝다. 그러나 분석을 위한 데이터의 형태는 하나의 제품에 대하여 전 공정에서 모든 측정값이 있기를 요구한다. 결국 상당량의 결측치가 존재하지 않을 수 없다. 어쩌면 이 요인이 LCD공정에서 마이닝을 함에 있어서 가장 큰 어려움일 수 있다. 이를 해결하기 위한 방안은 크게 두 가지가 있다. 하나는 분석 시에 사용하는 변수의 수를 가급적 줄이는 방법이다. 여러 가지 통계기법들을 이용해서 일차적으로 관련이 없는 불필요한 변수들을 분석에서 제외시킨 다음 분석을 하면 결측치의 양을 줄일 수

있다. 다른 하나는 각 제품별 결측치가 존재 할 때 타 제품과 비교 분석하여 유사한 제품의 값으로 대체하는 방법이다. 그러나 이 경우 결측치가 너무 많을 경우 그리 효용 있는 방법은 아니다.

### 2.2 결점/무결점의 비

최근 들어 결점을 줄이려는 시도 보다는 오히려 무결점을 최대의 과제로 공정을 관리하고 있다고 보아도 될 정도로 결점을 줄이기 위해서 부단한 노력을 기울이고 있고, 그 노력의 결과로 결점의 수가 상당량 줄어 있다. 그 결과 마이닝 분석에서의 가장 중요한 Target 변수인 결점여부를 나타내는 변수의 결점이 무결점에 비해 엄청나게 작은 비율로 발생한다. 이를 해결하기 위한 방법으로 Sampling 과정에서 결점과 무결점의 비를 조절하여야 한다. 결점과 무결점의 Sample의 비를 적절히 조절 한 다음 모델링 후 해석 시에 판정기준값(Cut Off Point)을 조절 하여 해석 하여야 한다.

### 2.3 데이터화 문제

데이터마이닝이란 대용량의 데이터를 효과적으로 분석하는 방법이다. 아무리 정확한 분석 틀이 있더라도 데이터가 존재하지 않는다면 데이터마이닝의 분석은 의미가 없어진다. 타겟 변수에 영향을 미치는 변수는 어떤 변수이고, 그 패턴이 어떤가에 따라서 향후 공정에 어떠한 조치를 줄것인가를 고려하여야 한다. 아무리 자동화 된 공정을 유지 하고 있더라도 타겟 변수에 영향을 많이 미치고 있는 변수들이 데이터화 되어 지지 않는다면 마이닝 분석의 의미는 그만큼 퇴색되어 질 것이다. 따라서 처음 공정을 설계하고 측정을 설계할 때 향후 마이닝 분석을 염두에 둔 설계가 이루어 져야 한다.

## 3. 품질마이닝 시스템 구축

앞서서도 언급하였다시피 품질 마이닝의 목적은 불량이나 결점에 영향을 미치는 요인을 찾는 것과 수율을 향상 시킬 수 있는 요인을 찾고자 하는데 있다. 이를 위해서는 일회성의 분석으로만 가능한 것이 아니라 지속적인 분석과 분석된 결과를 공정에 반영하고 반영된 이후의 변화를 감지하고 변화된 이후에도 더 좋은 공정을 만들기 위한 이러한 과정들을 반복함으로써 이루어진다. 이를 위해서는 품질 마이닝을 위한 하나의 시스템이 구축되어야 하고,

품질마이닝 시스템 구축을 위하여 갖추어야 할 몇 가지 사항들을 논하기로 한다.

### 3.1 시스템구축의 목적(필요성)

세계적으로 LCD산업의 시장 변화는 대량화, 무한 경쟁화, 급속한 기술발전등 무한 경쟁체계가 갖추어 졌다고 보아야 한다. 국내 뿐만아니라 국제적으로 그 시장의 규모가 커진것 뿐만 아니라 경쟁역시 치열한 상황이다. 이러한 무한 경쟁의 시장 상황에서 살아남거나 최고가 되기 위해서는 결국 품질 경쟁력에서 승리하는 방법이다. 이를 위해서는 효과적인 수율관리 체계와 무결점 공정 라인을 유지하는 것이라 하겠다. 기존의 SPC시스템에서 한발 나아가 수율 향상과 무결점 공정을 위한 체계화, 구체화, 자동화된 생산라인과 공정의 이상 발생 시보다 효율적인 관리 체계 구축을 위해서는 데이터마이닝 시스템의 구축이 필수일 것이다. 또한 다양한 품질/공정 관련 데이터의 효과적인 통합관리를 하고, 엔지니어가 발견하지 못했던 인자간의 패턴 밀유의차나, 공정 불량 의 원인 또는 중요인자들을 발견해서 이를 효율적으로 관리 할 수 있을 것이다.

### 3.2 주요내용

마이닝 시스템의 주요 내용을 살펴보면 다음과 같다.

#### 1) 주요변수 추출

반응변수(불량, 결점 등)에 영향을 미칠 수 있는 수많은 변수들 가운데 크게 영향을 미칠 수 있는 설명변수들을 선택해 내는 프로세스로서 반응변수의 종류에 따라 다양한 통계적 방법을 이용하게 된다.

#### 2) 유의차분석

반응변수에 영향을 미칠 수 있는 요소들의 복수 level에서 반응변수들의 값들이 통계적으로 의미 있는 차이를 나타내는가를 분석하고, 이에 기초하여 설비, 작업자 및 소재 등에 대하여 적절한 조치를 취하게 하며, 이를 통하여 대량의 부적합을 예방하는 효과를 기대 할 수 있다.(Joo et. al(2004), Devore(1995), Daniel(2000))

#### 3) 주요공정파악

반응변수에 영향을 크게 미치는 주요공정을 통계적으로 분석하여 해당공정의 중요도(기여도)를 평가한다. 이를 통하여 공정별 관리의 엄격 도에 대한

Plan을 수립/운영할 수 있다.

#### 4) 최적경로 탐색

주요공정변수들의 어떠한 조건에서 결점이 최소화 되고, 수율이 극대화 되는지를 탐색하는 기능으로서 recipe optimization을 주목적으로 이용한다.

#### 5) 예측 및 추정

여러 가지 환경 내지 조건들을 변화시켰을 때 수율 및 결점 수 등의 반응변수가 어떻게 변화하게 되었는지를 추정 또는 예측하는 내용으로 통계적 모델링을 통하여 반응변수의 추정/예측 값을 제공한다.

#### 6) 고장이나 결점 발생의 패턴인식

결점에 영향을 미치는 변수가 어떤 변수이고 어떤 level에서 혹은 어떤 조건에서 발생되어 진다는 식의 내용도 중요하지만 여러 가지 변수들이 조합적으로 어떤 환경에서 고장이나 결점이 발생하는 경향이 나타난다면 더욱 중요한 정보가 될 것이다. 이는 결국 전자의 경우가 OLAP차원의 분석이라면 후자의 경우는 데이터 마이닝을 통한 분석이라고 할 수 있을 것이다.

그 외 마이닝을 하기 위한 몇 가지 중요한 기능들로 구성되어 진다. 간단히 살펴보면 가장 기본적으로 데이터를 관리 가공, 정제할 수 있는 기능과 각종 도표나 그래프를 통하여 분석할 수 있는 기능 그리고 이러한 통계적인 분석 기능들을 이용하여 분석한 결과를 이용해서 report할 수 있는 기능들로 구성되어 있어야 할 것이다.

### 3.3 품질마이닝 시스템의 구성

마이닝 시스템의 구성은 데이터를 수집해서 마이닝을 할 수 있는 데이터를 제공해주는 DB부분, 각종 간단한 통계와 각종 도표 그리고 그래프 등을 볼 수 있는 SPC부분 및 여러 가지 주제별로 마이닝 기법들을 이용하여 마이닝 분석을 할 수 있는 마이닝 부분으로 구성이 된다.

#### 1) DB부분

마이닝 시스템에서 가장 기본이 되고 중요한 부분일 수 있다. 필요한 데이터를 수집하고 마이닝에 필요한 데이터로 수정, 정제, 변환을 하는 작업도 이루어 져야하며, 분석결과 각 주제별로 유용한 정

보들이 빠짐없이 분석에 포함될 수 있도록 설계되어야 한다. 또한 분석의 결과에 신경을 쓰지 않고 너무 많은 변수들을 DB에 포함 시킬 경우 불필요한 변수들이 너무 많아 발생하는 결측값이 많아지는 부분과 분석 시 요구되는 분석 시간을 엄청나게 필요로 하기 때문에 효율적이 되지 못하는 경우가 많다. 따라서 반응변수에 영향을 많이 미치는 변수들로 효율적인 DB화가 중요한 작업이라고 할 수 있다.

**2) SPC 부분**

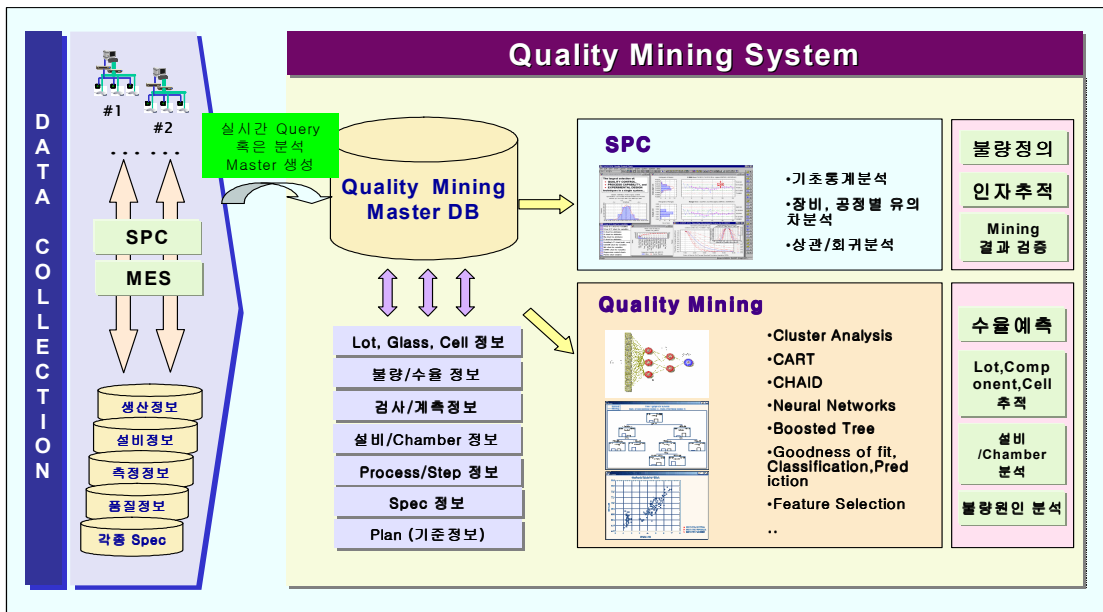
별도의 SPC시스템이 구축이 되어 있으면 이를 연동하여야 할 필요성이 있고, 그렇지 않을 경우 반드시 필요한 부분일 가능성이 있다. 마이닝 시스템에 가장 중요한 부분은 아니라고 할지라도 없어서는 안될 중요한 기능들을 포함하며, 전체 공정의 운영을 간략히 파악하기 위해서는 중요한 역할을 한다. SPC시스템은 대부분 통계패키지를 이용하여 구축이 되고, 여러 가지 통계량을 활용하는 경우도 있지만 대부분 각종 그래프를 이용한 분석이 많은

이 많다, 따라서 마이닝 부분의 구축 시 반드시 선행되어야 할 사항은 통계학적인 관점에서의 분석 혹은 컨설팅이다. 해당 변수들을 연결해서 마이닝 결과를 도출할 수 있는 기능을 보유하는 것이 중요한 것이 아니라 어떤 주제에 어떤 옵션을 사용하여 최적의 조건을 만들어 공정에 효율화를 극대로 상승하는 것이 가장 큰 궁극적인 목적이 될 것이다. 따라서 마이닝에 필요한 모든 조건들을 검토해서 시스템에 반영하여야 하며, 필요시 각 주제별로 선행 분석을 통해서 필요한 절차별로 정리해서 시스템에 반영하는 것도 하나의 방법이 될 수 있을 것이다.

<그림 3.1>은 이러한 내용의 마이닝 시스템의 구성도를 제시하고 있다.

**4) 통계패키지**

품질마이닝 뿐만 아니라 CRM에서의 마이닝에서도 가장 중요한 부분은 단순 마이닝 시스템의 구축이 아니라 분석이다. 이를 위해서는 좀더 신뢰성이 있



<그림 3.1> 품질마이닝 시스템의 구성도

부분을 차지하기 때문에 그래픽에 탁월한 기능을 갖고 있는 통계패키지의 선택이 중요한 사안일 수 있다.

**3) 마이닝 부분**

주제별로 필요한 변수들이 달라질 수 있고, 여러 가지 상황별로 선택되어지는 옵션이 변화될 가능성

고, 사용하기 편한 통계패키지가 필수 불가결 하다. 현재 이런 시스템을 구축하기 위하여 가장 많이 사용되고 있는 통계패키지로는 SAS, SPSS, STATISTICA, S-Plus등이 있다. 모든 패키지들이 장단점이 있지만, 단순한 통계량을 제시해 주는 기능보다는 데이터들을 보다 효율적으로 이해하기 위해서는 그래픽 면에서 강점이 있고 마이닝 분석을

초보자도 쉽게 활용할 수 있도록 응용이 편하게 되는 통계패키지의 선택이 좋을 듯하다.

STATISTICA를 채택하여 각 주제에 맞는 각종 분석을 쉽게 구현 할 수 있도록 구축되었다. (STATISTICA reference manual (2004))

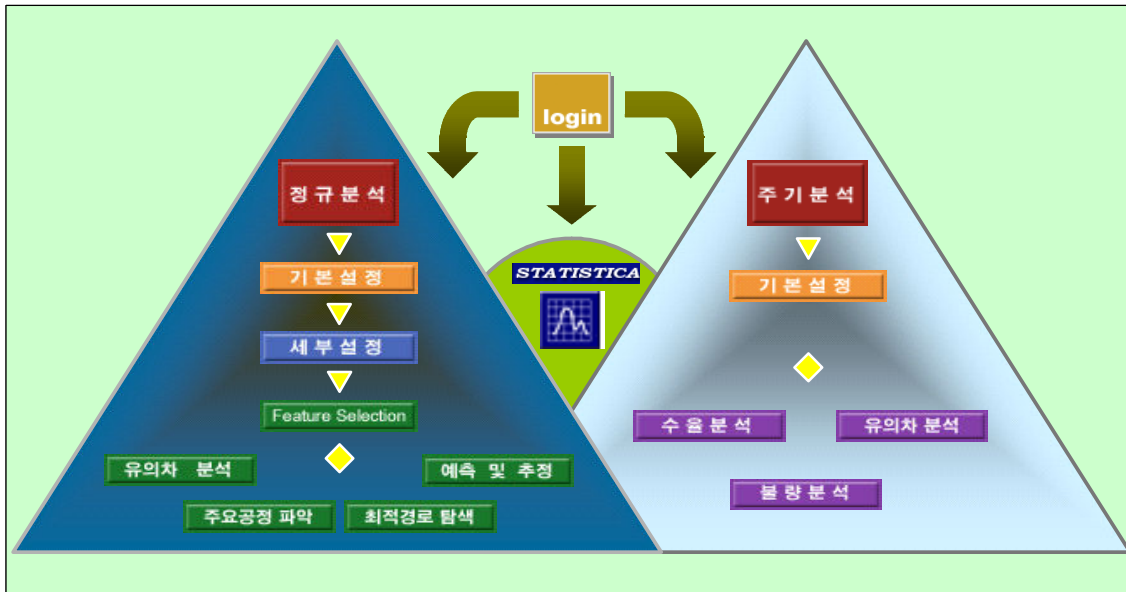
**4. 구축사례**

A사의 경우 품질에 영향을 미치는 인자들을 발견/관리 하고, 패턴을 발견하여 최적의 수율을 올리기 위해 마이닝 시스템을 구축하였다.

<그림 4.1>은 A사의 마이닝 시스템의 주요 내용들

**5. 결론**

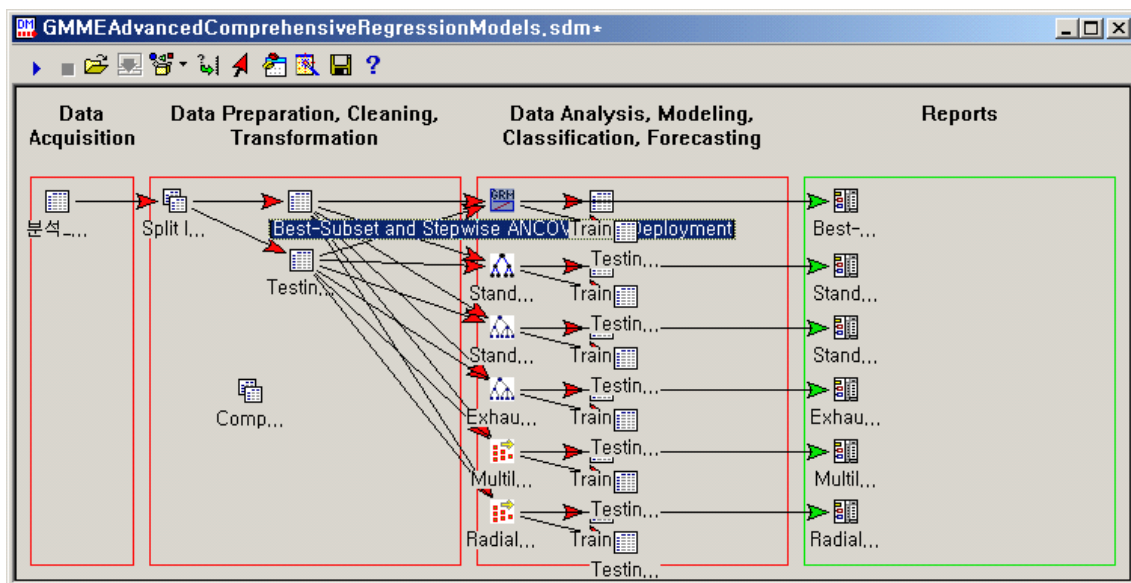
공정관리를 위해서 이제까지 SPC시스템에 만족해



<그림 4.1> A사의 품질마이닝 시스템의 주요 내용

을 도시하였다. A사의 경우 마이닝에 필요한 통계 패키지를 그래픽에 탁월한 기능을 보유하고, 데이터마이닝에 편리하고 우수한 기능을 보유한

왔던 것이 사실이다. 그러나 이는 단편적인 지식과 정보만을 제공해준다. 따라서 좀 더 복잡적이고 종합적인 정보를 통해서 좀 더 정확하고 효율적인 공



<그림 4.2> STATISTICA Data Mining 사례

2005 한국경영과학회/대한산업공학회 춘계공동학술대회  
2005년 5월 13일 ~ 14일, 충북대학교

정관리와 수율향상을 위해서는 데이터마이닝 시스템 구축은 필수불가결 하다고 하겠다.

CRM 시스템에서도 중요한 부분 중에 하나가 이런 마이닝 부분이 있지만 현재 대부분의 CRM시스템이 그리 만족스러운 효과를 달성하지 못한 이유는 전체 시스템의 효율과 분석적인 측면을 고려하지 않고 단순히 시스템 구축에만 모든 노력을 기울였기 때문이라고 생각한다. 이에 품질마이닝 시스템에서도 이를 교훈으로 삼아 좀 더 효율적인 시스템이 되기 위해서 통계적인 마인드와 이를 효율적으로 지원해 줄 수 있는 통계패키지의 선택으로 단순 시스템 구축에 그치지 않고 정확한 사용을 그리고 효율적인 관리를 할 수 있는 시스템 구축이 되어야 한다고 판단된다.

## 참고문헌

- Daniel, W.W. (1990). Applied Nonparametric Statistics, PWS-KENT Publishing Company
- Jay L. Devore (1995). Probability and Statistics for Engineering and the Sciences. Duxbury Press, 1995.
- H.S., Nam and Yulius, H. (2004). Feature Selection Methodology in Quality Data Mining, IEMS Spring Conference, SB7-5~8.
- STATISTICA System Reference(2004). StatSoft Inc.