

## SOFM신경망과 C4.5를 활용한 강의품질 개선

# Improving Lecture Quality using SOFM neural network and C4.5

이 장 희\*

한국기술교육대학교 산업경영학부

Jang-hee Lee\*

School of Industrial Management, KOREATECH, Cheonan 330-708, Korea

### [요약]

대학, 기업 및 학원에서 제공하는 교육 서비스의 질을 향상하기 위해서는 주요 활동인 강의의 품질 개선이 필수적이다. 강의 수행 후 수강생에 의해서 평가되는 강의평가 설문 데이터는 강의 품질을 측정하고 개선할 수 있는 좋은 도구로서, 대부분 간단한 통계분석을 통해 처리되고 있다. 본 연구는 강의평가 설문 데이터를 SOFM (Self-Organizing Feature Map) 신경망과 C4.5와 같은 분석도구를 사용하여 분석함으로써 수강생의 만족도와 강의 성과 관련한 특징을 보다 정확하게 파악하고 개선이 필요한 강의 품질 요소를 구체적으로 도출하여 강의 품질을 효율적으로 개선할 수 있는 방안을 제시하였다. 본 연구에서 제시한 방안을 국내 기업의 사내 강의에 적용한 결과, 만족도와 강의 성과 관점에서 미흡한 3개의 수강생 그룹에서 개선이 필요한 총 강의시간, 강의 자료, 강의 시간표 구성 요소를 개선하여 강의 품질이 향상되는 것을 확인하였다.

### [Abstract]

Improving lecture quality is very necessary for the service quality of education in universities, enterprises and education institutes. The student lecture evaluation survey data is a good tool for measuring lecture quality and have been often analyzed by simple statistical methods. This study presents an intelligent lecture quality improvement method that can improve student's overall satisfaction and performance by analyzing student lecture evaluation survey data. The method uses SOFM (Self-Organizing Feature Map) neural network and C4.5 to find the patterns in student's satisfaction and performance more correctly and then decide what to change in the lecture for the improvement of student's satisfaction and performance. We apply the proposed method to an enterprise lecture in Korea. We can find that it can improve the quality of an enterprise lecture by changing total lecture time, lecture material and organization of lecture schedule to be necessary improvements.

**Key Words:** C4.5, Education Service, Lecture Quality, Satisfaction, SOFM

<http://dx.doi.org/10.14702/JPEE.2014.071>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Received** 3 November 2014; **Revised** 12 November 2014

**Accepted** 18 November 2014

**\*Corresponding Author**

E-mail: [janghlee@koreatech.ac.kr](mailto:janghlee@koreatech.ac.kr)

## I. 서론

2000년도를 전후하여 많은 대학에서 강의평가제를 도입하여 실시함으로써 대학 교육의 주요활동인 강의에 대한 질 향상을 도모하고 있다. 현재는 대학 뿐 아니라, 기업과 학원에서 진행되는 많은 교육 프로그램의 강의에서 대부분 강의 종료 후에 강의평가를 실시하고 있다. 강의평가는 학생 또는 수강생들이 수강한 강의에 대하여 강의 만족도, 교수 및 강사의 강의 능력 등을 평가하고, 평가 데이터를 분석하여 강의방식과 내용을 개선함으로써 향후 양질의 강의를 제공하게 하는 효과를 발생시킨다. 이러한 강의평가의 중요성을 인식하여 강의평가 설문 문항을 합리적으로 구성하려는 연구가 진행되었다[1,2].

그러나, 실제 교육현장에서 진행되는 강의평가에서 평가 데이터를 분석할 때 간단한 통계분석이 수행되어 강의 품질 개선에 도움이 될 수 있는 의미있는 정보를 추출하지 못하고 있다. 대다수의 교육 현장에서는 강의평가 설문항목별로 평가 데이터를 단순히 비율 또는 평균 처리하여 분석하고 있다. 이러한 분석방법은 단순히 강의평가 항목별 전체적인 평균적 경향만을 확인할 수 있고, 구체적이고 정밀한 강의 개선 방향을 제시하지 못한다. 또한, 전체 수강생들에게 존재하고 있는 다양한 평가 특징들을 정밀하게 분석하지 못하고, 따라서 강의평가를 통한 수강생들의 정확한 요구 사항을 충분히 반영하지 못한다. 이에 본 연구에서는 강의평가 설문 데이터를 SOFM (Self-Organizing Feature Map) 신경망과 C4.5와 같은 지능형 분석도구를 사용하여 분석함으로써, 수강생의 만족도 및 강의 후 성과 관련한 특징을 정확하게 규명하고 개선이 필요한 강의 품질 요소를 구체적이고 명확하게 도출할 수 있는 방안을 제시하고자 한다.

논문의 구성은 다음과 같다. II장에서 강의품질과 강의평가 관련한 기존연구와 본 연구에서 활용한 분석 도구인 SOFM 신경망과 C4.5에 대해 살펴보았다. III장에서 SOFM 신경망과 C4.5를 활용한 강의품질 개선 방안과 이를 국내기업에서 진행된 강의에 적용한 사례를 설명하였다. IV장에서 본 연구의 결론에 대해 논하였다.

## II. 관련연구

### A. 강의 품질과 강의평가

강의 품질은 대학, 기업 및 학원과 같은 교육 현장에서 수행되는 강의의 질을 의미하는 일종의 교육 분야에서의 서비

스 품질이다. 교육의 질적 향상을 위해서는 양질의 강의가 제공되어야 하고 이러한 측면에서 대학에서의 강의평가는 강의 서비스 품질을 측정하는 도구로서 중요한 연구 주제였다.

Andrews는 강의평가가 강의의 수준을 향상하고 면학 분위기를 조성하며 이상적인 수업의 모습과 지향점을 제시한다고 하였다[3]. Peterson은 강의평가는 강의 목표 달성도, 의사전달, 교수와 학생간의 관계 형성, 교수법 등에 있어서 객관적이고 구체적인 지식을 학생과 교수에게 제공한다고 하였다[4]. 이기훈은 강의평가의 타당성과 신뢰성을 측정하는 방법을 제시하고 실제 자료에 적용하여 측정하였다[5]. 최경호와 이승주는 강의평가를 서비스 품질 향상의 측면에서 검토하여 KSI-SQI (한국서비스품질지수)에 기초한 강의 품질 평가 문항을 개발하였다[6]. 이외에도 대학 강의평가 문항의 신뢰성[7], 대학 강의 평가에 영향을 주는 요인[8] 등에 대한 선행연구가 진행되었다.

### B. SOFM 신경망과 C4.5 분석

본 연구에서는 강의 평가 설문 데이터를 정밀하게 분석하기 위해 SOFM (Self-Organizing Feature Map) 신경망과 C4.5 도구를 사용하였다. SOFM은 신경회로망의 일종으로 Kohonen에 의해 일반화되었다[9]. SOFM 신경망은 입력 속성의 수와 동일한 개수의 노드(node)로 구성된 입력 층과 보통 정방형의 노드로 구성된 출력층으로 구성된다. SOFM은 다변량의 입력자료에 존재하는 의미있는 패턴 또는 특징을 학습을 통해 추출하고, 비슷한 패턴을 가지는 입력 자료들을 출력 층의 특정 노드에 맵핑시켜 입력자료들을 그룹핑한다[9]. 즉, 추출된 다변량의 입력자료들의 패턴을 기준으로 입력자료들의 그룹을 구성하여 모든 입력자료에게 그룹을 할당한다.

C4.5는 Quinlan에 의해 개발된 대표적인 귀납적 학습 도구로서 학습 방법인 ID3 (Interactive Dichotomizer 3)를 변형한 의사결정 트리 기반의 분류자(decision tree-based classifier)이다[10]. C4.5는 어떤 개념에 관한 긍정적인 예와 반례로써 학습 데이터 집합을 구성하여 이로부터 개념을 구별할 수 있는 그림 1과 같은 의사결정 트리 형태의 분류규칙(Classification Rule)을 생성한다.

그림 1의 의사결정 트리는 10개의 입력속성과 3개의 클래스로 구성된 학습 데이터에 C4.5를 적용하여 생성한 경우이다. 의사결정 트리에서는 분류하고자 하는 출력변수의 개념들을 클래스(class)라고 하고, 클래스는 한정된 수의 입력속성의 조건으로 묘사된다. 생성된 의사결정 트리로부터, '만약 입력속성의 조건을 만족하면 출력 클래스가 나온다'는 분

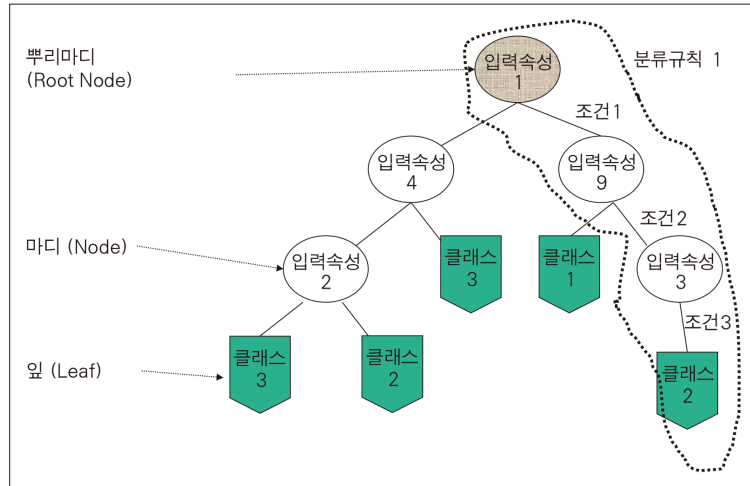


그림 1. C4.5 적용을 통해 생성된 의사결정 트리의 예  
 Fig. 1. Example of decision tree generated by C4.5 application.

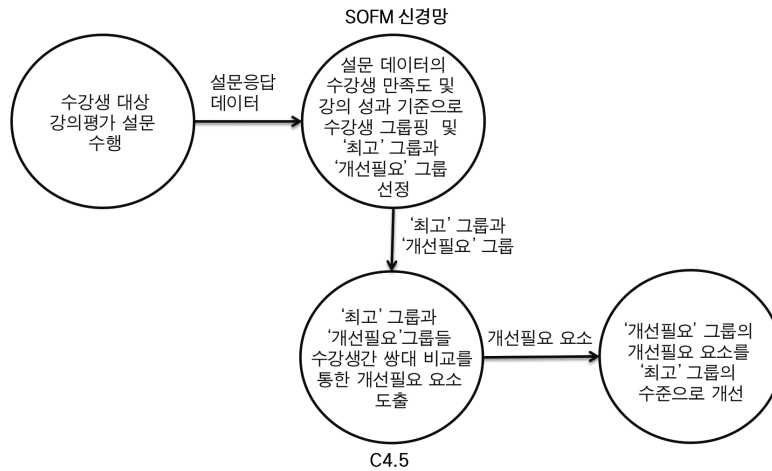


그림 2. SOFM신경망과 C4.5를 활용한 강의품질 개선 방안  
 Fig. 2. Lecture quality improvement method using SOFM neural network and C4.5.

류 규칙을 추출할 수 있다. 또한, 의사결정 트리 구조를 통해 출력 클래스에 대한 입력속성간 중요도를 파악할 수 있다. 의사결정 트리의 노드에 있는 입력속성들이 출력 클래스에 영향을 주는 중요한 속성이고, 특히 의사결정 트리의 뿌리 노드 (root node)에 있는 입력속성이 가장 영향력이 큰 중요한 속성이다. 따라서, <그림 1>의 의사결정 트리에 나타난 5개의 입력속성 1, 4, 9, 2, 3은 다른 입력속성보다 클래스의 변동에 많은 영향을 주는 중요한 속성이고, 그중에서도 뿌리마디에 있는 입력속성 1이 제일 중요한 속성이다. 다음으로 입력속성 4와 9가 중요하고 그 다음으로 입력속성 2와 3이 중요하다.

### III. SOFM신경망과 C4.5를 활용한 강의품질 개선 방안과 적용사례

본 연구에서 제안하는 SOFM신경망과 C4.5 분석에 기반한 강의품질 개선 방안은 그림 2와 같다. 그림 2에서 제시한 절차에 의해 국내의 '가' 기업에서 진행된 강의에 적용한 사례를 설명하고자 한다.

#### A. 수강생 대상 강의평가 설문 수행

강의 종료 후 수강생을 대상으로 강의평가 설문조사를 수

행하는 것이 본 연구에서 제안한 강의품질 개선 방안의 첫 단계이다. 일반적으로 강의평가 설문은 온라인 또는 오프라인상으로 진행될 수 있고, 설문 문항에는 강의 관련한 다양한 내용으로 구성할 수 있다. 본 연구에서는 합리적인 강의 서비스 품질 평가를 위해, 한국 서비스품질 지수에 기초하여 강의 서비스 품질을 측정할 수 있는 평가 문항들과 전반적인 강의 만족도 및 강의 후 얻은 강의 성과 관련한 문항 등을 구성하여 설문을 수행할 것을 권장한다.

본 연구에서 제안한 강의품질 개선 방법을 적용한 대상 강의는 기업 내부 엔지니어를 대상으로 3일간 진행된 강의로서, 이론과 실습 형태로 강의하고 수강생을 대상으로 서면으로 강의평가를 진행하였다. 강의에 참여했던 45명의 참가자를 대상으로 강의 자료, 총 강의시간, 강의 시간표 구성, 강의 진행 진도, 강의 내용, 수강생에 대한 강사응대, 강의 Skill, 강사 지식, 강사 실습지도, 강의장 시설, 교육 운영자의 지원, 전반적 만족도, 지식 성취도, 기술숙련 성취도와 같은 14개 항목으로 설문조사를 실시하였다. 각 설문 항목은 대부분 5점 척도로, 매우 불만족/매우 낮음 (1점), 불만족/낮음 (2점), 보통/중간(3점), 만족/ 높음 (4점), 매우 만족/매우 높음 (5점)으로 평가되었다.

**B. 강의 만족도와 강의 성과 기준으로 수강생 그룹핑 및 ‘최고’ 와 ‘개선필요’ 그룹 선정**

본 연구에서 제안한 강의품질 개선 방안의 두 번째 단계에서는 강의평가 설문 응답 데이터 중 강의 만족도와 강의 후 수강생이 느끼는 지식이나 기술 숙련의 성취도와 같은 성과 데이터를 기준으로 SOFM (Self-Organizing Feature Map) 신경망을 적용하여 수강생들을 그룹핑(grouping)한다. SOFM 분석결과 생성되는 수강생 그룹들 각각에 대해 해당 그룹에 속하는 수강생들의 만족도와 성취도 값의 특성을 비교하여 각 그룹의 특성을 파악한다. 수강생 그룹 중 가장 좋은 만족

총 강의시간 (<= 2 : 수강생그룹 4  
 총 강의시간 > 2  
     강의자료 (<= 4 : 수강생그룹 4  
     강의자료 > 4  
         강의시간표 구성 (<= 4 : 수강생그룹 4  
         강의시간표 구성 > 4 : 수강생그룹 2

**그림 3.** 강의평가 설문 데이터에 대한 C4.5 적용 결과: 수강생 그룹 4의 ‘개선필요’ 요소 도출

**Fig. 3.** C4.5 application result for lecture evaluation survey data: ‘need of improvement’ factor extraction for student cluster 4.

도와 성과를 가지는 그룹을 ‘최고’ 그룹으로 선정하고 나머지 그룹들을 ‘개선필요’ 그룹으로 분류한다.

본 연구에서 제안한 강의품질 개선 방법을 적용한 대상 강의에서 전반적 만족도, 지식 성취도, 기술 숙련 성취도에 대한 설문 응답 데이터를 기준으로 SOFM 신경망 적용을 통해 표 1과 같은 수강생 그룹을 구성하였다. 본 연구에서는 설문 응답자가 전체 45명으로 참가자가 적고 분석의 복잡성을 피하기 위해, SOFM 신경망을 통해 생성되는 클러스터 갯수를 4로 조정하였다.

4개의 수강생 그룹에는 각각 8명, 7명, 12명, 18명의 수강생이 속해 있고, 각 그룹에 속한 수강생들의 전반적 만족도, 지식 성취도, 기술 숙련 성취도의 평균 값으로부터 4 그룹의 특징을 파악할 수 있었다. 즉, 수강생 그룹 2는 만족도와 지식 및 기술 성취도가 다른 그룹과 비교해 보았을 때 월등히 높고 수강생 그룹 4는 다른 그룹들보다 매우 낮은 만족도와 지식 및 기술 성취도를 가지고 있고 따라서 강의 품질의 개선이 우선적으로 필요하다는 것을 확인할 수 있었다. 4개 그룹의 만족도, 지식 성취도, 기술 숙련 성취도 평균 값으로부터 그룹 ‘2’가 ‘최고’ 그룹이고 나머지 3개 그룹이 ‘개선필요’ 그룹에 해당함을 규명하였다.

**표 1.** SOFM 신경망 적용을 통한 확인된 수강생 그룹의 특성과 ‘최고’ 및 ‘개선필요’ 그룹

**Table 1.** Student clusters characteristics and ‘best’ and ‘need of improvement’ clusters generated by SOFM application

수강생 그룹	전반적 만족도 평균	강의성과		그룹 구분
		지식 성취도 평균	기술 숙련 성취도 평균	
그룹 1	3.98	3.72	3.67	‘개선필요’ 그룹
그룹 2	4.32	3.91	3.92	‘최고’ 그룹
그룹 3	3.33	3.18	3.18	‘개선필요’ 그룹
그룹 4	3.08	3.23	3.02	‘개선필요’ 그룹

### C. '개선필요' 그룹별 강의 품질 향상을 위한 개선필요 요소 추출

본 연구에서 제안한 강의품질 개선 방안의 세 번째 단계에서는 '최고' 그룹을 제외하고 모든 '개선필요' 그룹은 '최고' 그룹과의 쌍대 비교(pairwise comparison)를 통해 두 그룹에 속하는 수강생간에 서로 다르게 평가한 차별적인 강의 평가 설문항목을 C4.5을 사용하여 추출한다. 두 그룹의 강의 만족도와 성과 관련 설문 응답 데이터를 제외하고 모든 강의 품질관련 요소의 설문 응답 데이터에 C4.5 분석을 실시한다. C4.5 분석 결과로 생성된 의사결정 트리의 노드에 있는 강의 평가 요소들이 두 그룹을 분류하는 차별적 요소, 즉 개선이 필요한 요소이고 노드 중에서 의사결정 트리의 뿌리 노드에 있는 강의평가 요소가 두 그룹을 구분하는데 가장 큰 영향력을 가지는 중요한 요소이다. 일반적으로, '개선필요' 그룹이 다수 존재하므로, '개선필요' 그룹별로 '최고' 그룹과의 쌍대 비교를 통해 개선필요 요소를 각각 도출한다.

본 연구에서 제안한 강의 품질 개선 방법을 적용한 대상 강의에서 '개선필요' 그룹 중 가장 낮은 만족도와 성과를 가진 수강생 그룹 4에 대해 개선필요 요소를 추출하는 과정을 설명하면 다음과 같다. '최고' 그룹인 수강생 그룹 2와 '개선필요' 그룹인 수강생 그룹 4에 해당하는 수강생들의 강의 자료, 총 강의시간, 강의 시간표 구성, 강의 진행 진도, 강의 내용, 수강생에 대한 강사응대, 강의 Skill, 강사 지식, 강사 실습지도, 강의장 시설, 교육 운영자의 지원이라는 설문항목 응답 데이터를 C4.5 분석을 수행하여 그림 3과 같은 의사결정 트리를 생성하였다.

그림 3의 의사결정 트리로부터, 총 강의시간, 강의 자료, 강의 시간표 구성이라는 3가지 요소가 '개선필요' 수강생 그룹 4의 개선필요 요소로 확인되었다. 그 중에서도 생성된 의사결정 트리의 뿌리노드에 있는 총 강의시간이 가장 큰 요소로 판명되었다. 다른 '개선필요' 그룹에 대해서도 유사한 방식으로 개선필요 요소를 추출하였다.

### D. '개선필요' 그룹별 개선필요 요소 관점의 개선을 통한 강의 품질 향상

본 연구에서 제안한 강의품질 개선 방안의 마지막 단계에서는 '개선필요' 그룹 수강생들의 만족도 및 성과 향상을 위해 개선필요 요소 관점에서 개선한다. '최고' 그룹과의 비교 분석을 통해 도출된 개선필요 요소 관점에서 강의를 개선한다면 '개선필요' 그룹 수강생들은 그들의 만족도 및 성과 개선을 '최고' 그룹의 수준으로 유도할 수 있을 것이다.

수강생 그룹 4의 만족도 및 지식과 기술 숙련의 성취도 향상을 위해, 총 강의시간, 강의 자료, 강의 시간표 구성의 관점에서 개선 방향을 결정한 후 개선 작업을 수행하였다. 우선 현재 3일로 구성된 총 강의시간을 4일로 조정하였고, 강의 내용이 간단히 설명되어 있는 파워포인트 슬라이드로 구성된 강의자료에 더욱 자세한 내용 설명을 추가하여 재구성하였고, 4일간 진행되는 강의내용의 우선 순위를 조정하여 강의 시간표를 변경하였다. 다른 '개선필요' 그룹에 대해서도 도출된 개선필요 요소 관점에서 동일한 방식으로 진행하여 개선하였다. 이러한 과정을 통해 강의 품질의 향상을 유도할 수 있었고, 실제로 다음에 진행된 동일 강의에 대한 수강생 만족도 및 지식과 기술 숙련의 성취도 평균이 각각 4.232와 4.4338로 산출되었다. 이것은 이전 강의에서의 수강생 만족도 및 지식과 기술 숙련의 성취도 평균보다 약 16%와 20% 향상된 수준이었다.

## IV. 결론

대학, 기업 및 학원에서 수행하는 교육 서비스 품질이 향상되기 위해서는 강의 품질이 향상되어야 하고 이를 위해서 강의품질의 지속적인 개선이 이루어져야 한다. 본 연구는 대학, 기업 및 학원에서 실시하는 강의평가 설문 데이터를 SOFM 신경망과 C4.5와 같은 지능적인 분석도구를 사용하여 수강생의 만족도 및 성과 관련한 특징을 정밀하게 분석함으로써 강의 품질 향상을 위한 개선 요소를 체계적으로 도출하여 선택과 집중을 통해 효율적으로 개선할 수 있는 방안을 제시했다. 본 연구에서 제안한 방법을 기업에서 운영하고 있는 교육 프로그램의 강의에 적용한 결과, 동일 강의에 대한 수강생 만족도 및 강의 후 성취도의 평균 값이 이전 강의에서의 평균보다 각각 16%와 20% 향상되는 것을 확인할 수 있었다.

## 참고 문헌

- [1] H. L. Park, "A study on the improvement of a lecture evaluation tool in higher education -A case of improvement of a lecture evaluation questionnaire in "A" university-," *Journal of the Korea Academia-Industrial cooperation Society*, vol. 13, no. 11, pp. 5033-5043, 2012.
- [2] P. D. Cohen, "Effectiveness of student rating feedback for improving college instruction: A meta-analysis of find-



- ings,” *Research in Higher Education*, vol. 13, no. 4, pp. 321–341, 1980.
- [3] H. A. Andrews, *Teachers can be Fired! The Quest for Quality*. Peru, IL: Catfeet Press, 1995.
- [4] K. D. Peterson, *Teacher Evaluation: A Comprehensive Guide to New Directions and Practices*. Oaks, CA: Corwin Press, 1995.
- [5] K. H. Lee, “A study on validity and reliability of students’ evaluation,” *Journal of Korean Data and Information Science Society*, vol. 21, no. 1, pp. 87–98, 2010.
- [6] K. H. Choi and S. J. Lee, “A suggestion on instruction service quality assessment tool,” *Applied Statistics Research*, vol. 18, no. 2, pp. 487-497, 2005.
- [7] J. T. Kim and J. M. Lee, “A study on reliability of lecture evaluation by students,” *Journal of Korean Data and Information Science Society*, vol. 15, no. 1, pp. 183–191, 2004.
- [8] C. H. Ryu and J. H. Lee, “A study on student factors associated with the student evaluation of teaching at universities,” *Korean Management Review*, vol. 32, no. 3, pp. 789-807, 2003.
- [9] T. Kohonen, *Self-Organization and Associative Memory*, 3rd ed. Berlin: Springer-Verlag, 1989.
- [10] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA: Morgan Kaufmann Publishers, 1993.



**이 장 희 (Jang-hee Lee)**\_중신회원

1992년 2월 : KAIST 산업공학 석사  
1992년 3월 ~ 2002년 4월 : 삼성전자 반도체 사업부  
2001년 8월 : KAIST 산업공학 박사  
2002년 4월 ~ 현재 : 한국기술교육대학교 산업경영학부 교수  
<관심분야> 데이터 마이닝, 경영정보시스템, 품질 및 생산 공학, 기술경영 및 기술예측