

# 교육시설의 개념단계 공사비에측을 위한 인공신경망모델 개발에 관한 연구

## A Study on the Model of Artificial Neural Network for Construction Cost Estimation of Educational Facilities at Conceptual Stage

김 청 융\* · 손 재 호\*\*

Kim, Chung-Yung · Son, Jae-Ho

### 요 약

본 연구는 신축 교육시설 프로젝트의 개념단계에서 공사비를 예측하기 위한 인공신경망모델의 제안을 목적으로 한다. 현행 공공 교육시설의 개념단계 공사비에측에는 기본인자인 연면적에 의한 단일변수 모델이 적용되고 있다. 그러나 개념단계에서 단일변수 공사비에측모델을 적용하여 예측된 공사비는 그 오차범위가 크고, 실시설계 완료 후 물량산출에 의해 산정된 상세공사비와 비교하여 큰 차이를 보일 경우 프로젝트의 수정이 불가피하며, 이는 프로젝트의 비용을 증가시키고 공기를 지연시킨다. 그러므로 본 연구에서는 교육시설 프로젝트의 사업계획 수립 및 예산확보 과정에서 공사비에측에 적용이 가능한 다 변수 인공신경망모델을 제안하였다. 개발된 모델을 평가한 결과 평균오차율이 6.82%로써, 평균 93.18%의 정확도를 기록하였다. 제안된 인공신경망모델은 지난 5년간 신축된 교육시설의 공사예정금액을 실적자료로 사용하여 학습되었기 때문에, 차후 교육시설 신축공사의 예산편성에 그 활용이 기대된다.

키워드 : 공사비에측, 인공신경망모델, 교육시설 공사비

## 1. 서론

### 1.1 연구의 배경 및 목적

정부는 2005년 '연도별 BTL(Build Transfer Lease)추진 계획'에서 2007년까지 전국에 549개 초·중·고등학교의 신설계획을 발표하였으며, 2006년에는 당해 전국 157개교의 신설계획을 발표했다. 이는 선진국과의 학급당 학생 수(average class size) 비교에서 현저히 열악한 국내 교육환경을 개선시키기 위한 방안으로 판단되며, 차후에도 정부의 학급당 학생 수를 낮추기 위한 지속적인 노력이 예상된다. 또한 국내의 초·중·고등학교 시설 중 30년 이상의 노후교사 비율이 전체교사의 17%인 것으로 파악되고 있고, "향후 인구 증가율의 감소로 인하여 학생 수 증가가 둔화 되더라도 국내의 열악한 교육환경을 감안할 때, 교

육시설의 신·증축 필요성은 매우 높다"(권오현, 최민수 2004).

한편 프로젝트 전(全) 단계에서의 체계적인 사업비관리에 대한 관심이 고조되고 있는 가운데, 프로젝트 개념단계(conceptual stage)에서의 공사비에측(construction cost estimate)은 프로젝트의 성패와 관련하여 그 중요성이 대두되고 있으며, 많은 연구가 진행되고 있다. 일반적으로 개념단계에서의 공사비에측에는 기능단위, 연면적, 설비용량 등과 같은 기본인자(basic criteria)에 의한 단일변수(single-parameter) 공사비에측모델이 사용되고 있으며, 그 오차범위는 기관이나 연구자별로 차이는 있으나, "AACE(American Association of Cost Engineers)에 따르면 오차범위가 -30%~ +50% 정도로 그 신뢰성이 크게 떨어진다"(고성석 외 2005). 설계도면과 시방서가 작성된 이후 물량산출서(quantity takeoff)에 의한 상세공사비(detailed cost)가 산정되었을 때, 그것이 개념단계에서 예측한 공사비와 큰 오차를 가질 경우, 발주자는 사업변경에 따른 공기 지연 및 제반비용증가 등의 리스크를 책임지게 되거나, 심지어 사업을 포기하는 경우까지 발생한다.

현행 교육인적자원부의 사업진행 과정 중 개념단계 공사비에측에는 산정된 시설규모에 정부예산편성지침의 예산편성단가를

\* 일반회원, 서울시립대학교 건축공학과 석사과정(교신저자), uniyungs@hanmail.net

\*\* 일반회원, 홍익대학교 건축공학부 교수, 공학박사, jjwoodad@yahoo.co.kr

적용하고 있으나, 이는 단위면적당 단가(square meter prices)를 적용한 단일변수 공사비에측도모델로써 공사비에 영향을 미치는 다양한 내·외부적인 요인들을 고려하기 힘들고, 급변하는 정부의 교육정책을 반영하지 못하고 있다.

본 연구는 설계도서 작성 이후 물량 산출서에 의해 산정된 상세공사비에 공사비지수(construction cost index)를 적용한 다변수(multi-parameter)의 실적자료(historical data)를 인공신경망(Artificial Neural Network, 이하 ANN)에 적용하여 교육시설 프로젝트의 개념단계에서 공사비에측이 가능한 공사비모델을 제시하고, 그 효용성 검증은 목적으로 한다. 제시된 공사비에측모델은 차후 그 수요가 증가할 것으로 예상되는 교육시설 신축 프로젝트의 사업계획 수립 및 예산편성 과정에서 효율적이고 합리적인 의사결정을 지원하고, 단일변수 공사비에측모델을 적용하는 현재의 예산편성체계를 개선할 것이라 기대한다.

1.2 연구의 방법 및 절차

본 연구에서 공사비에측 ANN모델을 구축하기 위해 2001년에서 2005년까지 조달청 나라장터 홈페이지(www.g2b.go.kr)를 통해 설계시공분리방식(Design-Bid-Build)으로 발주된 전국 145개 신축 초·중·고등학교의 공사비를 실적자료로 수집하였다. 실적자료는 과거시점의 가치이므로 현재의 가치(present value)로 갱신(updating)하기 위해 공사비지수를 적용하고 이상 값을 제거하여 실적자료를 정상화하였다. 정상화된 실적자료에서 입력변수와 출력변수를 도출하였으며, ANN모델의 구축에는 Neuro Dimension, Inc.(www.nd.com)의 Neuro Solutions 5.0 evaluation ver.을 사용하였다. Neuro Solutions 5.0은 Microsoft의 Excel과 병행 사용가능한 Neuro Solutions for Excel을 제공하는데, 이는 ANN모델의 구축 및 그 결과 값 확인과정이 Excel Sheet상에서 이루어지므로 사용자가 보다 쉽게 접근 가능하다.

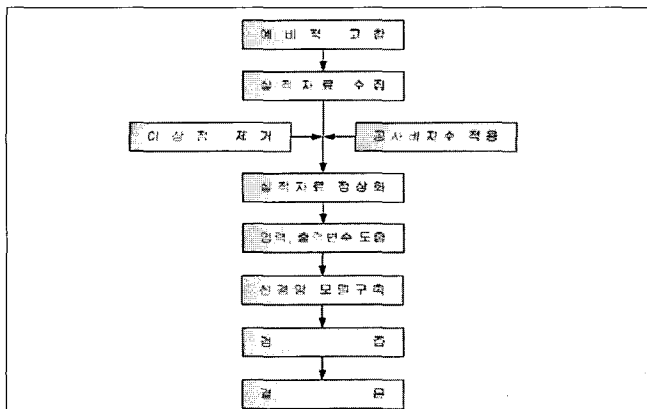


그림 1. 연구의 흐름도

본 연구에서는 총 64개의 ANN모델을 구축하고, 검증과정을 통해 최적의 모델을 선정하도록 한다. 그림 1은 본 연구의 과정을 도식화 한 것이다.

2. 선행연구 고찰

건설 분야에서 ANN을 적용한 연구는 다양하고 활발하게 진행되고 있다. 그 적용분야 중 하나는 공법선정에 ANN모델을 적용한 것이다. 김재엽 외 2인(2003)이 흙막이 지보공공법 선정모델에 ANN을 적용하였으며, 천봉호 외 3인(2004)이 지하말뚝공법 선정에 이를 적용하였다.

공사비에측분야에도 ANN을 적용한 연구가 활발히 진행되고 있으며, 그 예측 결과는 사례기반추론(Case-based Reasoning, CBR), 전문가시스템(Expert System), 회귀분석(Regression Analysis), 시물레이션(Simulation) 등의 방법과 비교하여 그 우수성이 입증 되고 있다. 김광희 외 1인(2004)은 ‘유전자 알고리즘에 의한 신경망 구조의 최적화를 이용한 공동주택의 초기 공사비에측에 관한 연구’에서 공사비에측에 있어 회귀분석모델보다 ANN모델이 더욱 정확한 예측결과를 보임을 입증하였다. 표 1은 국내 공사비에측분야에서 ANN모델을 적용한 대표적인 연구를 정리한 것이다.

표 1. 국내 인공신경망 적용연구

저자	연구시기	적용대상	연구 내용
김광희 외2인	2004.5	공동주택	공동주택 공사비 예측 정확도 비교
김선국 외2인	2000.9	사무소	신경망을 이용한 사무소 건물의 코스트 모델

3. 이론적 고찰

3.1 공사비 지수와 실적자료

실적자료는 과거시점의 가치이기 때문에 현재시점이나 미래 시점에 그 자료를 활용하기 위해서는 해당 시점의 자료로 갱신해야 한다. 이 과정에서 공사비지수를 적용하면 간편하게 과거 시점의 자료를 원하는 시점의 자료로 전환 가능하다. 본 연구에서는 자료의 정상화를 위하여 한국건설기술연구원에서 발표하는 건설공사비지수를 적용하였다.

건설공사비지수는 “건설공사에 투입되는 재료, 노무, 장비 등의 직접공사비를 대상으로 한국은행의 산업 연관표와 생산자물가지수, 대한건설협회의 공사부문 시중노임자료 등을 이용하여 작성된 가공통계로 건설공사 직접공사비의 가격변동을 측정하는 지수”(한국건설기술연구원 2005)로 정의되며, 기준년도는 2000년도로 설정되어 있다. 본 연구에서는 과거의 실적자료를 2005년의 6월의 가치로 전환하였으며, 전환에는 식(1)을 사용하였다.

$$PV = HV \times \frac{PI}{HI} \quad (1)$$

PV(Present Value) : 현재가치    HV(Historical Value) : 실적가치  
 PI(Present Index) : 현재지수    HI(Historical Index) : 실적지수

표 2는 본 연구에서 적용한 연도별 6월 건설공사비지수를 정리한 것이다.

표 2. 건설공사비지수 동향

구분	연도별 6월 동향					
	2000년	2001년	2002년	2003년	2004년	2005년
공사비 지수	100	102.0	105.7	115.3	125.3	128.1

### 3.2 인공신경망

#### (1) 인공신경망의 개요

ANN이란 “인간이나 동물이 가지고 있는 생물학적인 뇌의 신경세포(neuron)를 모델화하여 인공적으로 지능을 만드는 것이다. 즉 인간의 뇌에 존재하는 생물학적 신경세포와 이들의 연결 관계를 단순화시키고 수학적으로 모델링하여, 인간의 두뇌가 나타내는 지능적 형태를 구현하는 것이다”(임영도 1999). 따라서, ANN은 예제나 환경으로부터 학습시킨 후 복잡하고 연산이 많은 문제를 해결하는 것을 그 목적으로 한다. 일반적으로 퍼지, A.I. 등 다른 지능도구들과 비교하여 학습능력이 뛰어나지만, 네트워크 내부에서 발생하는 일들을 정확하게 수식으로 표현하지 않고 블랙박스(black box)로 간주하기 때문에 수학적 모델링 성능은 떨어진다.

#### (2) 인공신경망의 분류

ANN은 학습방식 및 그 구조에 따라 여러 가지 종류로 분류할 수 있다. 먼저 학습방법에 따라 지도학습(supervised learning) 방식과 자율학습(unsupervised learning)방식으로 분류할 수 있다. 지도학습방식은 사용자가 목표 값(desired output)을 지정하고 출력 값(actual network output)이 이 기준 값을 만족하

도록 학습하는 방식이고, 자율학습방식은 일정한 목표 값 없이 ANN이 알아서 출력을 내보내는 학습방식이다.

지도학습을 하는 ANN을 그 구조에 따라 두 가지로 분류하면, 입력층(input layer)과 출력층(output layer)만으로 연결되는 구조를 갖는 단층퍼셉트론과 입력층과 출력층 뿐만 아니라 하나 이상의 은닉층(hidden layer)을 가지는 다층퍼셉트론<sup>1)</sup>이 있다.

또한 다층퍼셉트론을 구조에 따라 두 가지로 분류하면, 피드포워드(feedforward) 네트워크와 피드백(feedback) 네트워크로 분류할 수 있다. 피드포워드 네트워크의 경우 입력신호가 은닉층과 출력층을 거쳐 피드백과정 없이 출력되지만, 리커런트(recurrent) 네트워크라고도 불리는 피드백 네트워크는 은닉층이나 출력층의 출력신호가 입력층으로 귀환되어져서 다시 출력되는 구조이다. 이러한 차이에 따라 피드포워드 네트워크를 정적 네트워크라 하고 피드백 네트워크를 동적 네트워크라 하기도 하며, 후자의 경우 동적 특성을 나타내는 장점이 있지만 연결가중치의 수가 증가하는 단점이 있다. 그림 2는 지도학습 ANN을 구조에 따라 분류하여 도식화 한 것이다.

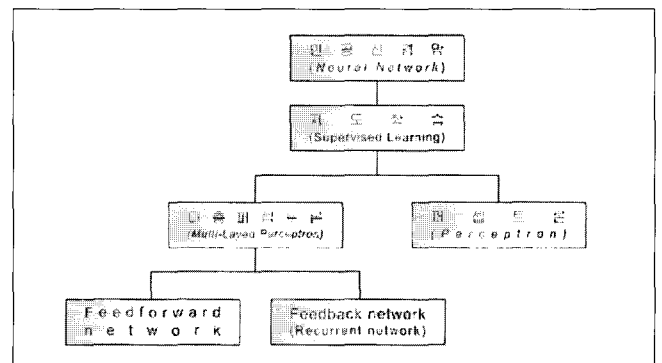


그림 2. ANN의 분류

본 연구에서는 수집된 실적자료의 공사비를 목표 값으로 설정하기 때문에, 지도학습을 하는 다층퍼셉트론 중 Feedforward 네트워크를 적용하였다.

#### (3) 인공신경망의 구조

그림 3은 다층퍼셉트론의 Feedforward 네트워크를 도식화한 것이다. 다층퍼셉트론은 입력층과 출력층 사이에 하나 이상의 중간층(은닉층)이 존재하는 구조이다. 신호는 입력층, 은닉층, 출력층 방향으로 전달되며, 은닉층과 출력층에서 입력층으로의 직접적인 연결이 존재하지 않음을 알 수 있다.

1) 다층퍼셉트론은 단순한 비선형 그룹핑이 불가능한 다층퍼셉트론의 문제점을 해결함으로써 ANN의 표준으로 사용되고 있으며, 그 응용범위가 다양하고 성능 또한 우수하다.

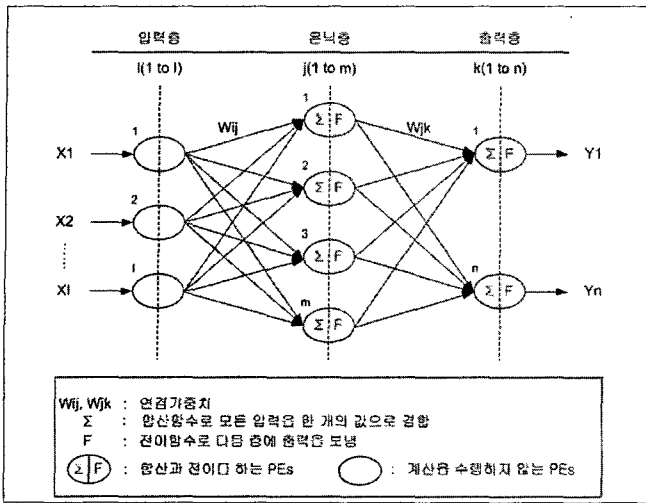


그림 3. ANN의 구조(김광희, 강경인 2004)

입력층은 외부로부터 데이터를 받는 PEs(Processing Elements, 처리요소)로 구성되고, 은닉층의 PEs에 정보를 보낸다. 출력층의 PEs는 은닉층의 PEs로부터 받은 입력 값을 계산하여 그 값을 외부로 보낸다. 은닉층과 출력층의 PEs는 합산함수와 전이함수를 가지고 있으며, 입력층의 PEs는 합산함수와 전이함수를 포함하지 않는다. 그리고 각 PEs 사이에는 연결가중치(weight)가 있다.

(4) 인공지능망의 학습방법

ANN의 학습이란 각 PEs와 PEs를 연결하는 적절한 가중치의 값을 찾아내어 원하는 동작을 하도록 만드는 과정을 말한다. 본 연구에서 적용하게 될 다층퍼셉트론의 학습에는 오류역전파(Error Back - Propagation) 알고리즘이 사용된다. 오류역전파 알고리즘은 델타법칙(Delta Rule)에 기초하고 있는데, 델타법칙은 “사용자가 지정한 목표 값과 학습에 의한 출력 값의 차이에서 만들어진 에러 값을 이용하여, 에러 값을 줄여나가는 방향으로 학습을 하는 법칙”(정승 2004)으로 정의 된다.

최초 무작위 선택에 의한 초기연결가중치(initial weight)가 결정된 후 학습자료(training data)가 입력이 되면, 초기연결가중치에 의한 순방향 계산이 이루어진다. 이후 순방향계산에 의한 출력 값과 입력자료의 목표 값의 오차가 계산되고, 오차에 의한 역방향 계산이 이루어져 연결가중치가 수정된다. 수정된 연결가중치에 의해 모든 자료의 MSE(Mean Square Error)<sup>2)</sup>가 계산되고, 그 값이 학습정지조건을 만족할 경우 학습은 종료되고, 만족하지 않을 경우 학습자료 입력단계로 되돌아가 학습을 반복한다.

퍼셉트론 ANN은 오차를 만드는 과정에서 전이함수인 비선형 함수 값을 통과한 값이 취해진다. “대표적인 비선형함수로는 지수함수(Exponential), 선형함수(Linear function), 시그모이드 함수(Sigmoid function), 가 있다”(Linoff, G. and Berry 1997). 이 중 그림 4의 시그모이드함수는 대표적인 미분가능 함수로써, 본 연구에서 은닉층과 출력층의 전이함수는 기존연구에서 그 우수성이 입증된 시그모이드함수를 사용한다.

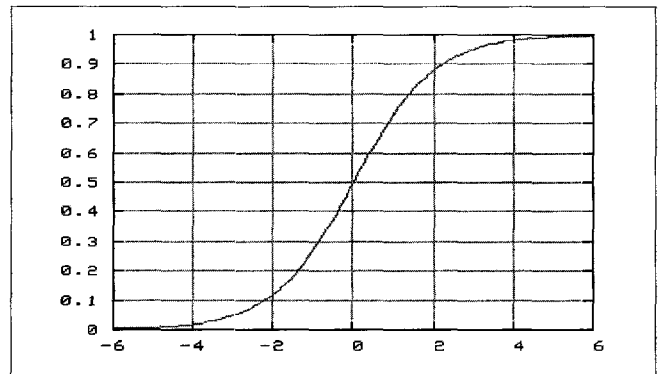


그림 4. 시그모이드함수

4. 인공지능망 모델의 구축

4.1 실적자료의 수집

ANN모델을 통한 개념단계 공사비예측을 위해 2001년부터 2005년까지 조달청 나라장터 홈페이지를 통해 공고된 145개 교육시설 신축프로젝트에 대한 공사예정금액을 실적자료로 수집하였다. 공사예정금액은 예정가격<sup>3)</sup>에 부가가치세와 관급자설치 관급액이 포함된 금액이다. 수집된 공사예정금액은 건설공사비 지수를 적용하여 2005년 6월의 가치로 갱신하였으며, 그림 5~그림 7은 수집된 실적자료의 지역별, 교육단계(level of

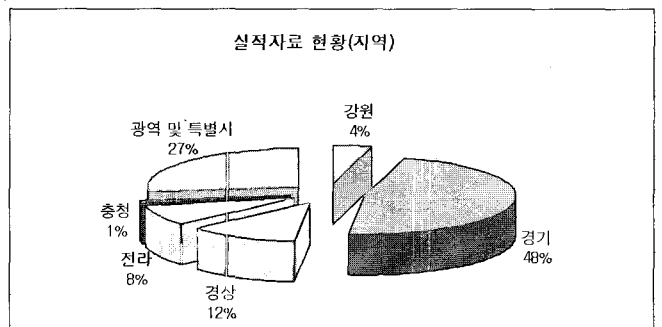


그림 5. 지역별 실적자료현황

2) 오차제곱평균,  $MSE = \frac{1}{n} \times \sum (\text{출력값} - \text{목표값})^2$

3) 원가계산에 의한 가격으로 재료비, 노무비, 경비, 일반관리비, 이윤을 포함한 금액

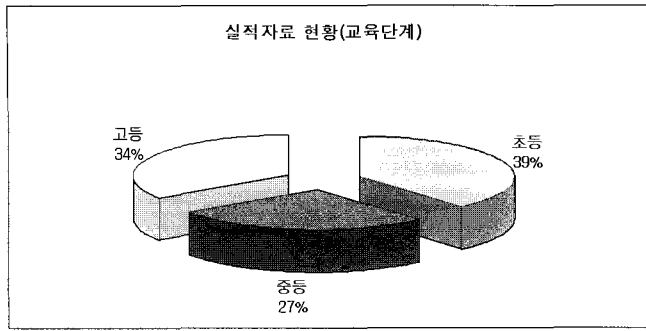


그림 6. 교육단계별 실적자료현황

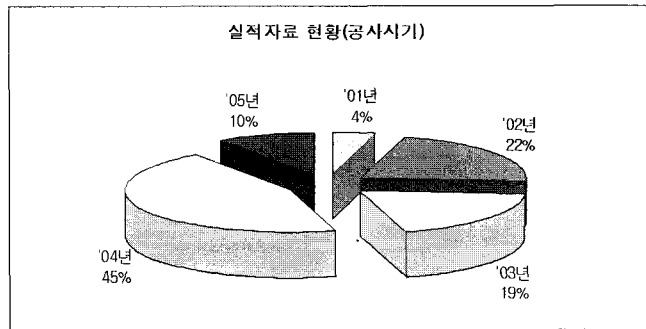


그림 7. 공사시기별 실적자료현황

education)별, 공사시기별 비율을 도식화 한 것이다. 지역별로는 경기도가 가장 많은 비율을 차지했고, 공사시기는 2004년도 가 가장 많은 비율을 차지했다.

4.2 입·출력변수의 설정

수집된 실적자료에서 ANN모델을 구축하기 위한 입력변수와 출력변수를 설정하였다. 본 연구에서는 문헌 및 기존 공사비에 측연구에서 공사비에 영향을 미치는 요소로 선정한 다양한 변수들 중 프로젝트 개념단계에서 수집 가능한 6개의 변수를 입력변수로 설정하였다. 입력변수와 출력변수의 단위 및 입력 값 코딩(coding)에 관한 자세한 내용을 표 3에 정리 하였다.

표 3. 입·출력 변수의 단위 및 입력값

구분	변수	단위	입력값
출력변수	공사비	십억원	양의 실수 (4.89~15.22)
입력변수	공사시기	-	양의 정수 (1~5) (2001년(1), 2002년(2), 2003년(3), 2004년(4), 2005년(5))
	지역	-	양의 정수 (1~6) (강원(1), 경기(2), 경상(3), 전라(4), 충청(5), 광역 및 특별시(6))
	교육단계	-	양의정수 (1~3) (초등(1), 중등(2), 고등(3))
	지하층수	-	양의 정수 (0~2)
	지상층수	-	양의 정수 (1~7)
	연면적	천㎡	양의 실수 (4.28~14.26)

출력변수로는 공사비(construction cost)를 설정하였으며, 이는 신축공사<sup>4)</sup>와 기계설비공사, 전기공사, 정보·통신공사, 소방공사 그리고 조정공사의 공사예정금액 합계이다. 입력변수의 설정에 있어 공사시기는 공사예정금액 중 금액비율이 가장 높은 신축공사가 공고된 시점을 기준으로 정리하였으며, 지역변수는 강원, 경기, 경상, 전라, 충청, 광역 및 특별시의 6개 지역으로 분류한 후 숫자로 전환하여 입력하였다. 또한 교육단계는 초등, 중등, 고등으로 구분 후 숫자로 전환하여 입력하였다.

4.3 제어변수의 설정

ANN모델의 설계에 있어 그 모델의 수는 제어변수의 설정에 따라 무수히 많고, 그 결과 값 역시 큰 차이를 보인다. 따라서 사용자는 적절하게 제어변수의 범위를 한정하여 다수의 모델을 구축하고 검증과정을 통해 최적의 모델을 선정한다. 제어변수로는 구조변수, 학습변수 및 학습정지 조건이 있으며, 이들은 실적자료와 더불어 ANN모델의 성능을 결정하는 중요한 요소이다.

(1) 구조변수

입력층과 출력층의 구조에서 각층의 수는 1개이고, 입력변수와 출력변수를 설정함으로써 각층의 PEs수는 결정된다. 은닉층에서 사용자가 설정해야하는 구조변수에는 은닉층의 수와, 각 은닉층의 PEs 수가 있다. 은닉층의 수가 증가함에 따라 각 PEs 사이의 연결가중치 수도 증가하므로 ANN의 계산시간은 증가하게 된다. 일반적으로 은닉층의 수는 1개에서 시작하여 그 수를 늘려간다. 본 연구에서 1개의 은닉층에 대해서 공사비예측 ANN 모델을 구축하였다. 은닉층의 PEs의 수는 출력층의 PEs의 수보다 많아야 하며, 이를 만족하는 조건에서 최대한 작은 수에서 시작하여 그 수를 증가시켜가면서 모델을 구축한다. 본 연구에서는 각 학습변수의 조합에 대해 은닉층의 PEs수를 2개에서 7개까지 변화시켜 가면서 모델을 구축하고, 그 결과 값을 검토한다.

(2) 학습변수

학습변수의 설정은 오류역전파 알고리즘의 매개변수(intermediate variable)로써, 은닉층과 출력층 각각의 step-size와 모멘텀(momentum)이 있다. step-size는 일반적으로 학습률(training rate)로 표현되기도 하며, “ANN모델의 학습변수 중 가장 민감한 변수이다”(정슬 2004). step-size가 지나치게 작으면 수렴하지 않는 경우가 발생하며, 지나치게 크게 설정

4) 건축공사, 토목공사, 기타 부대공사의 예정금액 합계

할 경우 ANN의 안정성이 떨어진다. 일반적으로 step-size는 0에서 1사이의 값을 취하며, 처음에는 작은 값에서 시작하여 수렴 속도를 빨리하기 위해 점차 늘려간다.

모멘텀은 이전 샘플에서 가중치변화율과 현 샘플에서의 가중치변화율 중 어느 것에 더 많은 비중을 두느냐를 결정짓는 변수이다. 이는 학습과정에서 발생하는 지역최소점(local minima)<sup>5)</sup>으로 오차가 수렴하는 것을 방지해 주는 역할을 한다.

본 연구에서는 학습변수의 범위를 한정하기 위해 출력층의 모멘텀 값을 0.7로 고정시키고 은닉층의 step-size와 모멘텀 그리고 출력층의 step-size를 변화시켜 가면서, 시행착오적인 방법으로 64개 모델을 구축하고, 그 결과를 검증하였다. 표 4는 ANN모델을 구축하기 위한 각 학습변수의 범위를 정리한 것이다.

표 4. 학습변수의 설정

구 분	학습변수	변 수 값
은닉층	step-size	$0.6 \leq \eta \leq 0.9$
	모멘텀	$0.6 \leq \alpha \leq 0.9$
출력층	step-size	$0.6 \leq \eta \leq 0.9$
	모멘텀	$\alpha = 0.7$

(3) 학습정지

ANN은 과잉학습으로 인해 성능이 저하되는 경향이 있다. 그림 8의 학습횟수에 따른 MSE변화에서, 최적 학습횟수를 넘어 학습횟수가 증가함에 따라 학습자료에 대한 MSE 값은 점점 감소하지만, 교차확인(cross validation)<sup>6)</sup>자료의 MSE 값은 증가함을 알 수 있다. 이는 ANN모델의 학습횟수가 무한히 증가하면, 학습 자료에만 일반화된 모델이 구축되기 때문이다.

따라서 본 연구에서는 두 가지 학습정지 조건을 설정하였다. 먼저 최대 학습횟수(maximum epochs)를 1000회로 한정하였다. 그리고 교차확인자료의 MSE 값 변화추이에 따라 학습을 정지시키기 위해 교차확인자료의 MSE가 100동안 향상이 없을 경우 학습을 정지하도록 설정하였다.

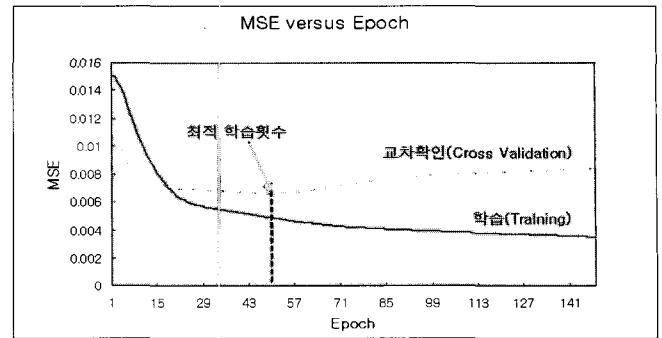


그림 8. 학습횟수와 ANN모델(Linoff, G., and M. J. A. 1997)

4.4 이상 값의 제거

이상 값(outlier)란 실적자료 중에서 대부분의 자료들과 아주 멀리 떨어진 개체들을 말한다. 그 원인으로는 최 극한 값을 실적자료로 수집, 자료의 코딩오류, 결측 값을 실제 값으로 인식 등이 있으며, 이와 같은 이상 값이 ANN의 학습과정에 포함될 경우 모델의 성능에 악영향을 주므로 적절한 조치가 요구된다. 본 연구에서는 이상 값 제거를 위해 제어변수를 달리한 9개의 예비 ANN모델을 구축하고 전체자료에 대한 목표 값과 출력 값의 표준화된 잔차(standardized residual)를 구하여, 그 값의 절대 값이 2이상인 실적자료를 이상 값으로 정의하고 10개의 실적자료를 제거하였다.

이상 값이 제거된 135개의 실적자료는 NeuroSolutions 5.0의 무작위 선택(random selection)방식을 사용하여 표 5와 같이 분류하였다. 전체자료는 시스템구축에 필요한 자료와 평가를 위한 자료로 구분하였으며, 시스템구축자료는 학습자료와 교차확인자료로 구분하였다.

표 5. 실적자료의 분류

구 분	수 량	
시스템구축	학 습	80
	교차확인	20
평 가	35	
합 계	135	

4.5 인공신경망 모델의 평가

학습자료와 교차확인자료의 입력변수와 출력변수를 ANN에 입력하고 각 제어변수를 조정하여 총 64개의 모델을 구축하였으며, 학습정지 후 35개 평가자료(testing data)를 입력하여 목표 값과 출력 값의 평균오차율을 표 6에 정리하였다. 표 6의 평균오차율은 각 제어변수 조합에 대해 3차례 학습을 실시한 후 최적

5) 학습과정에서 오차는 광역최소점(Global Minima)으로 수렴해야 하지만 시스템이 비선형이 강하고 복잡하면 지역최소점(Local Minima)에 빠져 더 이상 오차가 수렴하지 않는 경우가 발생하며, 비선형성이 강할수록 지역최소점의 수가 증가한다.

6) ANN모델이 과잉학습으로 인하여 학습자료에 지나치게 일반화되는 것을 방지하기위해 학습과정에서 모델의 일반화 수준을 결정하기 위해 사용된다.

모델에서의 평균오차율을 정리한 것이다<sup>9)</sup>.

표 6. 학습변수에 따른 평가자료 평균오차율(%)

은닉층		출력층 step-size			
step-size	모멘텀	0.6	0.7	0.8	0.9
0.6	0.6	7.49	7.91	7.85	6.98
	0.7	7.83	7.53	7.26	7.88
	0.8	8.22	7.02	7.91	7.58
	0.9	8.46	7.79	7.87	7.32
0.7	0.6	7.65	7.75	7.73	7.37
	0.7	7.89	7.70	7.04	7.55
	0.8	7.31	7.78	6.94	7.15
0.8	0.6	7.70	7.84	7.73	8.07
	0.7	7.63	7.73	7.43	7.03
	0.8	7.85	7.19	7.85	7.42
	0.9	7.78	8.47	7.60	8.68
0.9	0.6	8.06	8.02	7.65	7.68
	0.7	<b>6.82</b>	7.83	8.00	7.47
	0.8	7.91	7.39	7.50	7.87
	0.9	8.17	7.29	7.53	7.48

평가용 자료의 오차율과 평균오차율은 아래의 식을 적용하였다.

$$Error\ Rate(\%) = \frac{|DO - AO|}{DO} \times 100 \quad (2)$$

DO(Desire Output) : 목표 값, AO(Actual Network Output) : 출력 값

$$Average\ Error\ Rate(\%) = \left( \sum_{i=1}^n |Error_i| \right) / n \quad (3)$$

64개 모델 중 최저 평균오차율을 기록한 모델의 경우, 학습변

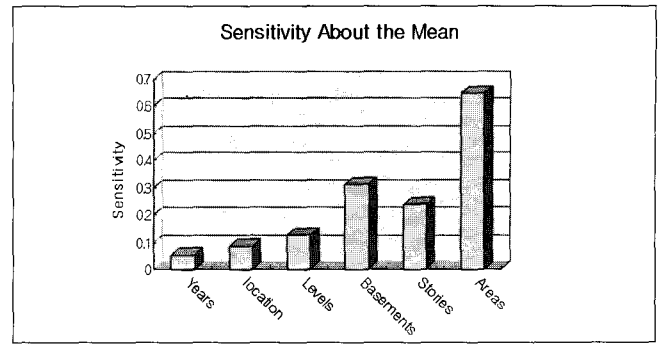


그림 9. 입력변수 민감도분석

수를 은닉층 step-size 0.9, 모멘텀 0.7, 출력층 step-size 0.6, 모멘텀 0.7로 설정하였을 때, 평가자료에 대한 평균오차율이 6.82%로써 평균 93.18%의 정확도를 기록하였다. 이 모델의 은닉층 PE는 7개이며, 학습자료의 학습횟수가 371회일 때 학습은 정지하였다. 이 모델을 본 연구에서 제안하는 최적 ANN모델로 선정하였으며, 35개 평가자료에 대한 오차율을 표 7에 정리하였다.

그림 9는 공사비에 대한 각 입력변수의 평균에 대한 민감도(sensitivity about the mean)<sup>8)</sup>를 도식화한 것이다. 입력변수 중 연면적이 0.64로 가장 높은 수치를 나타내었으며, 지하층수

7) ANN은 무작위 선택에 의해 초기연결가중치를 결정하므로 동일한 제어 변수에 대해 매 회 다른 결과 값을 나타낸다.

8) ANN모델의 테스트 과정에서 입력변수 사이의 상대적 중요성을 측정하고, 입력변수에 대해 모델의 출력변수가 어떻게 변화하는가를 표현하는 수치

표 7. ANN모델에 의한 공사비에측 결과

No.	목표 값 (십억원)	출력 값 (십억원)	오차율	No.	목표 값 (십억원)	출력 값 (십억원)	오차율	No.	목표 값 (십억원)	출력 값 (십억원)	오차율
1	12.28	11.61	5.44%	13	9.42	9.76	3.59%	25	8.36	9.20	10.02%
2	10.54	9.69	8.07%	14	10.06	10.06	0.05%	26	12.11	12.20	0.75%
3	10.09	11.31	12.12%	15	8.70	9.05	4.03%	27	11.51	10.20	11.41%
4	12.44	12.80	2.92%	16	10.36	11.79	13.75%	28	8.22	9.66	17.58%
5	10.96	10.35	5.52%	17	11.06	11.53	4.28%	29	11.01	9.54	13.40%
6	8.68	9.23	6.40%	18	9.08	9.01	0.76%	30	11.52	10.35	10.18%
7	9.34	9.20	1.47%	19	10.82	10.10	6.62%	31	10.80	11.98	10.95%
8	8.30	8.18	1.41%	20	10.72	10.52	1.86%	32	10.08	11.07	9.78%
9	9.83	9.62	2.17%	21	9.61	9.12	5.10%	33	10.73	11.23	4.66%
10	9.64	9.34	3.09%	22	10.28	10.54	2.53%	34	9.37	8.11	13.46%
11	11.91	11.34	4.78%	23	10.72	10.07	6.06%	35	8.30	7.76	6.51%
12	11.61	12.46	7.36%	24	7.44	5.91	20.63%	평균오차율		<b>6.82%</b>	

가 0.31로 두 번째로 높은 수치를 기록하였다. 그리고 공사시기는 0.05로써 본 연구를 통해 구축된 ANN모델의 입력변수 중 가장 작은 값을 기록하였다.

## 5. 결 론

본 연구에서는 교육시설 신축 프로젝트의 개념단계 공사비에 측에 있어, 기존의 모델을 보완하기 위해 2001년에서 2005년 사이에 신축된 전국 135개 초·중·고등학교의 실적자료를 이용한 ANN모델을 제안하였으며, 그 결과 평균오차를 6.82%인 공사비에측 모델을 구축하였다. 그리고 본 연구에서 제안한 ANN모델의 입력변수 민감도분석을 통해 입력변수로 설정한 모든 변수가 공사비에 영향을 미치는 것을 확인하였다. 이를 통해 제안된 ANN모델이 공사비에 영향을 미치는 다양한 내·외부적인 요인을 반영한 융통성 있는 모델로서 현행 단일변수 공사비에측모델의 한계를 보완해 줄 것이라 판단된다. 본 연구의 결과 및 차후 개선사항은 아래와 같다.

(1) ANN모델은 입·출력변수의 설정 및 학습변수의 설정에 따라 그 성능에 있어 큰 차이를 보인다. 일반적으로 “교육시설 공사에서 전체공사비에 대한 토목공사비의 비율은 전체공사비의 7%수준”(박태근 외2인 2004)인 점을 감안할 때, 본 연구에서는 지하층수를 입력변수로 설정하여 토공사에 대한 공사비를 일부 고려하였지만, 토공사에서 큰 비율을 차지하는 기초 및 지평공사에 대한 입력변수를 설정하지 않아 그에 대한 공사비를 고려하지 못했다. 차후 연구에서 기초 및 지평공사에 영향을 미치는 요소를 입력변수로 설정하여 ANN모델을 구축한다면 보다 높은 성과를 기대할 수 있을 것이라 판단된다.

(2) ANN모델은 실적자료에 의한 학습에 그 바탕을 두고 있기 때문에 우수한 성능의 모델을 구축하기위해 실적자료의 축적은 매우 중요하다. 이러한 관점에서 다른 실적자료에 비해 일반적이지 못한 이상 값이 학습자료에 포함될 경우 ANN모델의 성능에 부정적인 영향을 미치게 되므로, 실적자료의 축적에 있어 이상 값의 진단 및 제거를 위한 노력이 요구된다. 또한 정부의 교육정책은 해를 거듭할수록 학급당 학생 수를 줄이고 교육시설의 고급화를 추구하는 등 전반적인 교육시설 수준을 향상시키는 방향으로 추진되고 있다. 이로 인해 교육시설의 공사비는 매년 증가하고 있으며, 이러한 상황을 반영하는데 지나치게 오래된 실적자료 역시 우수한 ANN모델을 구축하는데 부정적인 영향을 준다. 본 연구에서 ANN모델의 학습과 평가에 135개의 자료를 사용한 것을 감안할 때, 150개 정도의 실적자료를 유지하는 수

준에서 신규 사업의 실적자료를 축적하여 갱신하고 일정 시간이 경과한 실적자료를 모델 구축 시 제외시킨다면, 보다 정확한 교육시설의 공사비에측을 위한 ANN모델을 구축할 수 있을 것으로 판단된다.

## 참고문헌

1. 권오현, 최민수(2004). 건설상품별 중·장기 시장 전망(Ⅱ) - 비주택부문-. 한국건설산업연구원.
2. 고성석 외(2005), 건설관리학, 사이텍미디어, pp. 120.
3. 김광희, 강경인(2003). 공동주택 프로젝트의 초기 공사비 예측을 위한 신경망 학습에 유전자 알고리즘을 적용한 모델에 관한 연구, 대한건축학회 논문집(구조계) 19권 10호, 대한건축학회, pp. 113-142.
4. 김광희, 강경인(2004). 유전자 알고리즘에 의한 신경망 구조의 최적화를 이용한 공동주택의 초기 공사비 예측에 관한 연구, 대한건축학회논문집 구조계 20권 2호, pp. 81-88.
5. 김광희, 김상용, 강경인(2004). 공동주택 공사비 예측 정확도 비교에 관한 연구, 대한건축학회 논문집(구조계) 20권 5호, 대한 건축학회, pp. 93-101.
6. 김두섭, 강남준(2000). 회귀분석 기초와 응용, (주)나남출판, pp. 52-75.
7. 김대수(1994.2). 신경망 이론과 응용(I), 하이테크정보 출판부.
8. 김선국, 구인환(2000). 신경망을 이용한 사무소 건물의 코스트 모델, 대한건축학회 논문집(구조계) 16권 9호, 대한 건축학회, pp. 59-66.
9. 김재엽, 서장우, 강경인(2003). 신경망을 이용한 흙막이 지보 공공법 선정모델 개발에 관한 연구, 대한건축학회 논문집(구조계) 19권 5호, 대한건축학회, pp. 121-128.
10. 박태근 외 2인(2004). 학교건축의 공사비 분석 및 예측에 관한 연구, 한국건설관리학회 학술발표대회 논문집 제5회, 건설관리학회, pp. 290-295.
11. 배상현, 백형래, 이진섭(1999). Neural Network의 이론과 응용, 조선대학교출판부.
12. 임영도, 이상부(1999). 퍼지·신경망·유전진화, 도서출판 영과일.
13. 정 슬(2004). 인공지능시스템 I 신경회로망의 구조 및 사용법, 충남대학교출판부.
14. 천봉호, 구충완, 엄익준, 구교진(2004). 건축공사 말뚝공법 선정을 위한 신경망모델 개발, 한국건설관리학회 학술발표대회 논문집 제5회, 건설관리학회, pp. 317- 322.
15. 한국건설기술연구원(2005). 2005년 6월 건설공사비지수 동향



16. Clifford, J. Schexnayder, Richard E. Mayo.(2004). Construction Management Fundamentals, McGraw-Hill, pp. 132-161.
17. Donald S. Barrie, Boyd C. Paulson(1992). Professional Construction Management: including C.M., Design-construct, and General Contracting, 3rd edition, McGraw-Hill, pp. 307-331.
18. Linoff, G. and Berry, M. J. A.(1997). Data mining techniques : For marketing, sales, and customer support, John Wiley & Sons, Inc., pp. 286-359.

논문제출일: 2006.04.19

심사완료일: 2006.06.05

---

### Abstract

The purpose of this study is propose an Artificial Neural Network(ANN) model for the construction estimate of the public educational facility at conceptual stage. The current method for the preliminary cost estimate of the public educational facility uses a single-parameter which is based on basic criteria such as a gross floor area. However, its accuracy is low due to the nature of the method. When the difference between the conceptual estimate and detailed estimate is huge, the project has to be modified to meet the established budget. Thus, the ANN model is developed by using multi-parameters in order to estimate the project budget cost more accurately. The result of the research shows 6.82% of the testing error rates when the developed model was tested. The error rates and the error range of the developed model are smaller than those of the general preliminary estimating model at conceptual stage. Since the proposed ANN model was trained using the detailed estimate information of the past 5 years' school construction data, it is expected to forecast the school project cost accurately

**Keywords** : Construction Cost Estimation, Artificial Neural Network, Construction Cost of Educational Facility

---