

인공신경망을 적용한 기획단계의 도로건설 공사비 예측 모델

Road Construction Cost Estimation Model in the Planning Phase Using Artificial Neural Network

한형동* · 김정환** · 윤중호*** · 서종원****

Han, Hyeong Dong · Kim, Jeong Hwan · Yoon, Jung Ho · Seo, Jong Won

Abstract

Construction cost estimation in planning phase which calculates the cost for performing construction tasks is used for various ways. Meanwhile, in the case of road construction, the existing cost estimating method in early phase based on numerical mean value of the past is not accurate to be used. This paper propose neural network model for estimating road construction cost in planning phase to solve the limit of current cost estimating method. The model was designed using past road construction bidding records, and variables of model were optimized through trial and error. The estimation result of the model was compared with regression analysis and government's standard and it was verified that the model is better in accuracy. It is expected that the proposed model will be used for road cost estimation in planning phase.

Keywords : *planning phase, cost estimation, road construction, artificial neural network, regression analysis*

요 지

공사비 견적이란 건설공사를 수행하는데 소요되는 공사비를 산정하는 것으로, 공사 기획단계에서 사용하는 개략견적은 다양한 목적으로 사용된다. 한편, 도로공사의 경우, 현행 사용되는 개략견적방법은 과거의 실적자료를 산술평균하여 단위 km 당 공사비를 제시한 것으로 실제 개략공사비로 활용하기에는 어려움이 있다. 본 연구에서는 기획단계의 도로공사견적 산정방식에 대한 한계를 해결하기 위하여, 인공신경망을 적용한 개략공사비 모델을 제시하는 것을 목적으로 한다. 도로건설 공사 입찰을 위한 기본설계도면 작성 이후 공사비를 결정하는 변수를 포함한 과거 실적 자료를 바탕으로 인공신경망을 적용하여 기획단계에서 공사비 예측이 가능한 도로 공사비를 개략견적 하는 모델을 설계하였으며, 시행착오방식을 이용하여 모델의 변수들을 최적화하였다. 또한 최적화된 모델을 통해 나온 결과를 회귀분석을 적용한 결과와 비교하여 정확도가 뛰어난 것을 검증하였다. 제안된 인공신경망모델은 차후 도로건설공사의 개략공사비 산정을 위한 활용될 것으로 기대된다.

핵심용어 : 기획단계, 공사비 예측, 도로건설, 인공신경망, 회귀분석

1. 서 론

1.1 연구의 배경 및 목적

공사비 견적이란 건설공사를 수행하는데 소요되는 공사비를 산정하는 것으로, 예산편성, 대안검토, 입찰, 공사수행 등 각 단계에 따른 의사결정과정의 중요한 근거가 된다(Oberlender와 Trost, 2001). 공사비 견적 방법은 공사시행 단계에 따라 가용한 정보의 수준이 달라지므로 각각 다른 방법으로 결정되며, 그 정확도에 따라 개략견적, 개산견적 및 상세견적으로 구분된다(박홍태, 2006).

이 중, 공사 기획단계에서 사용되는 개략견적은 해당 시설물에 대한 부분적인 정보 또는 과거 사례에 기초하여 공사

비를 예측하는 방법으로서, 발주자에게는 예산편성의 기준을 마련하여 주고, 설계자에게는 견적된 금액 내에서의 합리적 설계를 가능하게 해주며, 시공자에게는 실행 예산산정, 공사 규모 파악 등을 할 수 있게 해주는 등 다양한 역할을 한다(곽수남 등, 2009; 김선국 등, 2000).

하지만, 공사의 특성상 표준화된 공법 및 설계를 구축하기 힘든 토목공사의 경우 과거의 실적자료를 산술평균하여 개략견적에 이용하는 것은 한계가 있으며, 이러한 방식의 개략견적의 한계를 극복하기 위하여 과거부터 많은 연구가 선행되어져 왔다.

한편, 토목공사가 주를 이루는 도로공사의 경우, 국토해양부(2009)의 도로업무편람에서 개략공사비의 산정 기준을 제

*국토해양부 대전지방국토관리청 시설주사 · 한양대학교 공학대학원 건설관리학 석사 (E-mail : hanhd@korea.kr)

**정희원 · 한양대학교 대학원 건설환경공학과 박사과정 (E-mail : kimjh0418@gmail.com)

***주)대우건설 사원 · 한양대학교 대학원 건설환경공학과 석사 (E-mail : jung.ho.yoon@daewooenc.com)

****정희원 · 교신저자 · 한양대학교 건설환경공학과 부교수 · 공학박사 · P.E. (E-mail : jseo@hanyang.ac.kr)

시하고 활용하도록 하고 있지만, 이 기준도 역시 과거의 실적자료를 단순 산술평균하여 단위 km 당 공사비를 제시한 것으로 실제 개략공사비로 활용하기에는 어려움이 있다.

본 연구에서는 현재 도로공사건설에 대해 기획단계의 개략공사비 견적방식에 대한 한계를 해결하고 더 정확한 공사비 견적을 위하여, 1990년대 이후 초기 공사비 예측 방법으로 활발하게 연구가 진행된 인공신경망(김광희와 강경인, 2003) 기법을 활용하여 개략공사비 모델을 제시하는 것을 목적으로 한다. 도로건설 공사 입찰을 위한 설계도면 작성 이후 물량 산출서에 의해 산정된 상세 공사비 및 공사비를 결정하는 변수를 포함한 과거 실적 자료를 바탕으로 인공신경망을 적용하여 기획단계에서 공사비 예측이 가능한 도로 공사비를 개략견적 하는 모델을 제시할 것이며, 모델을 활용하여 예측한 결과를 다중회귀분석을 통해 나온 예측결과 및 현행 기준을 적용한 결과를 비교하여 신경망 모델의 정확도를 검증한다.

1.2 연구의 범위 및 방법

본 연구에서는 인공신경망이 적용된 도로 공사비 개략견적 모델을 설계하기 위하여 조달청 나라장터 홈페이지(www.g2b.go.kr)를 통해 발주된 도로 건설 공사를 수집하였다. 수집 대상은 개찰일 기준으로 2003년부터 2010년 1월 사이에 국토해양부 산하의 5개 지방국토관리청이 실 수요기관으로 발주된 일반 국도 신설공사로 한정하였으며, 총 93개 데이터를 수집하였다.

인공신경망모델 설계를 위하여 도로 총연장, 도로의 폭 등을 입력변수로 설정하고 기초금액을 출력변수로 설정하였으며, 인공신경망모델 설계 후 The MATLAB V.7.8 버전의 Neural Network Toolbox를 이용하여 모델을 구현 및 실험하였다. 또한, 설계한 모델을 최적화하기 위하여 인공신경망 모델 설계에서 주로 사용하는 시행착오(trial and error)방식을 이용하여 인공 신경망 내부의 변수들을 조정하였으며, 최적화된 모델에서 나온 예측 결과를 다중회귀분석(multiple regression analysis)을 적용한 결과 및 국토해양부의 기준을 적용한 결과와 비교하여 인공신경망 개략공사비 산정 모델에 대한 검증을 실시하였다.

2. 선행 연구 고찰

2.1 선행 연구 고찰

인공신경망모델은 분야에 대한 배경지식 없이 입/출력변수를 조정함으로써 공사추진 단계별로 공사비 예측을 쉽게 적용할 수 있는 장점과 별도의 분류 없이 정성적 데이터와 정량적 데이터를 함께 사용할 수 있는 장점(Bode, 1998)을 가지고 있어 과거부터 인공신경망을 이용한 공사비 예측 연구가 활발히 진행되어 왔다.

Hegazy와 Ayed(1998)는 도로공사비 예측에서 회귀분석이 가지는 한계점을 지적하고, 인공신경망을 이용한 공사비 예측 모델을 제시하였다. 이 때, 인공신경망의 각 뉴런을 연결하는 연결가중치를 결정하기 위해서 유전 알고리즘을 이용하였다. 또한, Wilmot과 Mei(2005)도 기존의 선형회귀분석을 이용한 예측 방법에 의한 공사비 예측의 한계를 지적하

고 인공신경망을 이용하여 도로공사의 공사비 예측모델을 제시하였다. 이때, 정확한 예측을 위하여, 공사에 투입되는 자재, 인력, 장비, 계약의 성격 및 계약 환경 등의 과거의 자료를 이용하여 인공신경망을 훈련시켜 공사비 예측 모델을 제시하였다. 국내에서 인공신경망을 이용하여 공사비 예측을 시도한 연구로는 김광희 등(2006)의 경우가 있다. 이 연구는 프로젝트 초기 단계에서의 공사비 예측에서 회귀 분석의 한계를 극복하여 정확하고 신속한 공사비 예측을 위하여 인공신경망을 이용하여 공동주택의 공사비를 예측하는 모형을 제시하였다. 이 때, 신경망의 연결가중치, 모멘텀 등의 알고리즘에 쓰이는 매개변수들을 찾는 데 시간을 최소화하며 정확성을 향상시키기 위해 유전자 알고리즘을 이용하였다. 또한, 김청용과 손재호(2006)는 신축 교육시설 프로젝트의 개념단계에서 공사비 예측의 중요성을 강조하고 현행 사용되는 예측 모델의 한계를 극복하기 위하여 인공신경망을 이용한 공사비 예측 모델을 제시하였다.

앞서 소개한 바와 같이 현재까지 인공신경망을 이용하여 공사비 예측 모델을 설계하고 내부 변수를 최적화하는 연구가 국내·외로 많이 진행되어왔다. 하지만, 국내에서 인공신경망을 이용하여 도로건설공사의 개략공사비를 예측한 연구는 현재까지 없었다.

한편, 인공신경망모델의 결과와 비교를 실시하는 다중회귀분석은 과거 수행자료를 바탕으로 공사비 예측모델을 도출하는 방식으로, 수학적 명확성과 데이터 분석의 용이함으로 공사비 예측 모델 구축을 위한 연구에 활발히 활용되어 왔다. 곽수남 등(2009)은 사업 초기단계를 크게 기획단계와 설계초기단계로 분류하여 각 대표공종을 독립변수로 설정하고 다중회귀분석의 단계별 선택법(stepwise method)을 활용한 개략공사비 산정 모델을 제시하였고, 윤유중과 우성권(2006)은 변수 간 상관관계 분석을 통해 토공, 구조물공, 포장공, 터널공의 공사비 증감비율을 독립변수로 선정하고 직접공사비 증감비율을 종속변수로 활용하여 고속도로 건설공사 예비비 산정을 위한 다중회귀분석 기반의 모델을 구축하였다. Petroutsatou 등(2006)은 도로 터널 공사비를 예측하기 위해 33개 터널에서 수집된 149개 자료를 바탕으로 2단계 다중회귀분석기법을 활용하여, 암반특성과 버팀대 수량의 관계를 파악하고, 추정된 수량을 기반으로 비용을 산출하였다.

2.2 현행 도로공사 단가산출 기준 고찰

현재 국내 도로공사의 경우, 국토해양부(2009)의 도로업무편람에 따라 일반도로건설에 대해 평균단가, 표준단가, 교량 및 터널 평균단가를 제시하고 있다(표 1, 표 2). 하지만, 제시된 단가는 '06-'08년 사이에 공사 발주된 도로사업 설계단가의 평균치로서 지방부/도시부로 나누어 확장/신설에 대해 단위길이상 단가를 제시하고 있으며, 특수구조물, 연약지반

표 1. 일반국도건설 평균단가(국토해양부, 2009)

(단위 : 억원/Km, 제비비 포함)

구 분	차로수	계	시설비	용지비
지방부 (국도, 확장)	2 → 4	145	128	17
도시부 (국대도, 신설)	4	223	178	45

표 2. 일반국도건설 표준단가(국토해양부 2009)
(제잡비 포함)

구 분	차로수	토공 및 기타	교 량	터 널
지방부 (국도, 확장)	2→4	86억원/km	489억원/km	287억원/km
도시부 (국대도, 신설)	4	134억원/km	467억원/km	259억원/km

과다지역 등에 대하여 한계를 보이는 등 도로건설의 다양한 공사적인 특성을 반영하지 못하고 있다(국토해양부, 2009).

3. 이론 고찰

3.1 인공신경망

본 논문에서 공사비 예측을 위해 사용될 인공신경망은 인간이 가지고 있는 생물학적 신경세포(neuron, 뉴런)를 수학적으로 모델링한 것으로 표본이 된 생물학적 뉴런은 그림 1 과 같이 크게 수상돌기(dendrite), 축색돌기(axon), 세포체(cell body, soma)로 이루어져 있다. 이 때, 각 뉴런은 인접 뉴런의 시냅스로부터 수상돌기를 통하여 들어온 신호가 임계값을 넘으면, 축색돌기로 신호를 내보내어 다른 신경으로 신호를 전달하는 과정으로 자극을 전달한다(한학용, 2009).

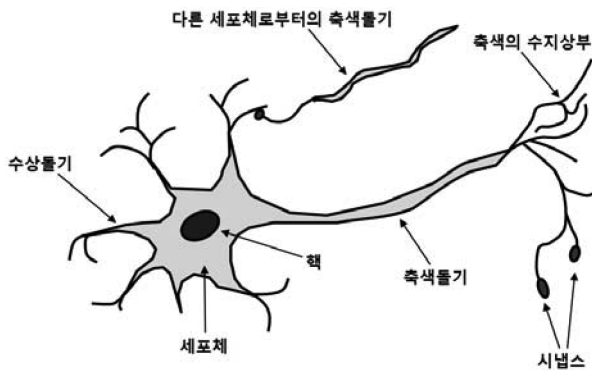


그림 1. 생물학적 신경세포(한학용, 2009)

인공신경망의 기본 구조도 그림 2와 같이 1) 일정 값을 입력받아 2) 뉴런 내부에서 처리를 한 후, 3) 출력하는 구조로 되어있다. 먼저 입력단계에서는 각각의 가중치(weight, W_i)와 곱하여진 입력값(X_i)과 활성화 함수의 임계값의 변동을 흡수하기 위한 가중치값인 편견값(bias, b)이 합산되어 뉴

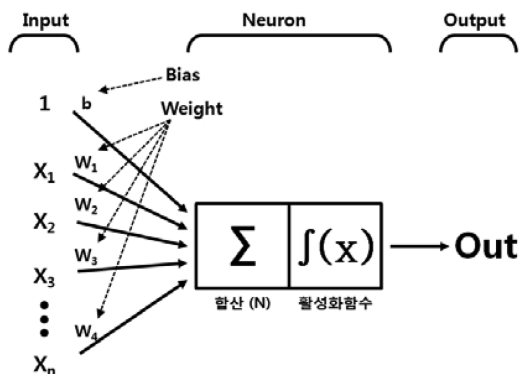


그림 2. 인공신경망의 기본 구조

런에 입력된다(식 (1)).

$$N = \sum X_i W_i + b = (X_1 W_1 + X_2 W_2 + \dots + X_n W_n) + b \quad (1)$$

합산되어 입력된 값(N)은 활성화 함수(activation function)를 거쳐 결과를 출력하는데, 사용되는 활성화 함수는 여러 종류가 있으며, 그 중 일반적으로 많이 사용되고 본 논문에서도 사용한 활성화 함수를 그림 3에 나타내었다.

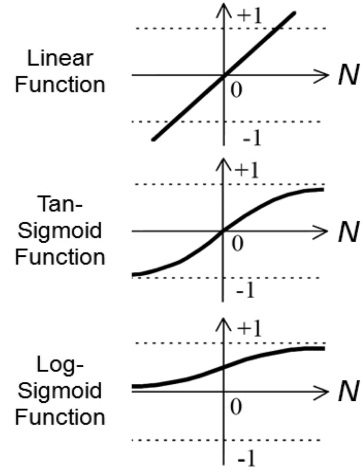


그림 3. 활성화 함수

인공신경망은 구조, 출력형태, 학습방법 등에 따라 여러 가지로 분류된다. 우선 인공신경망은 계층 수에 따라 단층(single-layer)신경망과 다층(multi-layer)신경망으로 나뉜다(그림 4). 단층신경망은 가장 단순한 구조로 입력층(input layer)과 출력층(output layer)으로 구성되어 있으며, 입력층은 외부에서 입력값을 받는 역할을 하고, 출력층은 신경망의 결과를 출력하는 역할을 한다. 다층신경망은 입력층, 출력층, 그리고 하나 이상의 은닉층(hidden layer)으로 구성되어있다. 은닉층은 직접적으로 결과를 출력하지는 않고 출력층 및 다음 은닉층으로 결과를 전송하는 역할을 하며, 신경망의 구조를 복잡하게 함으로써 복잡한 문제의 최적 해를 찾아내는데 유용하게 사용된다.

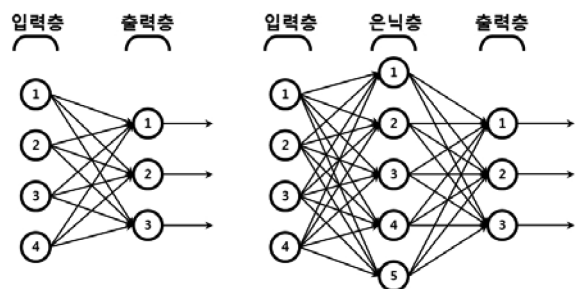


그림 4. 단층신경망(좌) 과 다층신경망(우)

그밖에, 구조에 따라 전방향(feed-forward)신경망과 회귀(feed-back 또는 recurrent)신경망으로 구분되고, 학습방법에 따라 지도학습(supervised learning), 자율학습(unsupervised learning), 경쟁식 학습(competitive learning)신경망으로 구분된다.

인공신경망모델에 사용되는 입/출력 샘플 자료는 크게 훈

련(traning)과정, 입증(validation)과정, 시험(test)과정으로 나뉘어 사용된다. 훈련과정은 인공신경망모델 초기상태에서 입력값에 대한 결과값이 목표값에 대하여 오차를 최소한으로 갖도록 뉴런을 연결하는 연결 가중치와 편견값을 변경시키는 과정을 말하며, 인공신경망을 주어진 샘플에 대해 알맞은 모델로 최적화시킨다. 이 때, 신경망모델이 훈련에 사용되는 샘플에만 과도하게 맞추어져 지역적으로 최적화가 되는 것을 과적합(over fitting) 현상이라 부르는데, 입증과정은 과적합 현상을 방지하기 위하여 매회 훈련이 종료된 시점에서 입증과정에 할당된 샘플을 사용하여 모델에 대한 오차를 계산함으로써 훈련의 지속/종료 여부를 판단한다. 마지막으로 시험과정은 훈련 종료 후 모델의 성능을 평가하는 역할을 한다.

3.1.1 역전파 알고리즘

인공신경망에서 가장 일반적으로 쓰이는 오류 역전파(back-propagation) 알고리즘은 인공신경망에서 역사적으로 가장 중요한 개발 중 하나(Hegazy 등, 1994)로 크게 순방향으로 입력값으로부터 목표값에 대해 결과값을 계산하는 단계와 결과값과 목표값 사이의 오차값으로부터 역방향으로 입력층까지 연결가중치를 갱신하는 단계로 나누어지며, 기준치 이하로 오차값이 나올 때까지 두 단계를 반복적으로 실시한다.

이 때, 연결가중치의 변화를 결정하는 변수는 크게 두 가지가 있는데, 하나는 학습률(learning rate), 하나는 모멘텀(momentum)으로 불린다. 학습률은 해당 층에 산출된 연결 강도의 변화량을 얼마만큼 적용시킬 것인지를 결정하며, 모멘텀은 바로 이전의 연결 강도의 변화량을 얼마만큼 적용시킬 것인지를 결정한다. 학습률이 너무 작으면 수렴하지 않는 경우가 발생하고, 너무 크면 인공신경망의 안정성이 떨어진다. 또한, 모멘텀은 지역 최소점으로 오차가 수렴하는 것을 방지하여 하여 주며, 학습 시간을 줄이는데 유용하게 사용된다(한학용, 2009).

3.2 회귀분석

회귀분석(regression analysis)이란 어떤 현상을 지배하는 변수사이의 관계를 설명하고 수학적 모형을 유도하는 통계 기법으로, 독립변수의 관측값으로부터 종속변수의 값을 예측하거나 추정하는 분석 기법이며, 독립변수의 개수가 하나일 경우 단순회귀분석(simple regression analysis)이라하고 두 개 이상일 경우 다중회귀분석이라 한다(김해경과 윤진희, 2004).

독립변수를 x 라 하고 종속변수를 y 라 할 때, 일반적으로 독립변수와 종속변수 간 관계를 식 (2)과 같이 표현한다.

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in} \quad (2)$$

여기서, α 는 절편을 의미하며 상수라 하고, β 는 회귀계수라 한다.

본 연구에서는 두 개 이상의 독립변수와 종속변수의 관계를 분석할 때 적합한 다중회귀분석을 활용하였으며, 분석은 독립변수들 간의 관계를 의미하는 다중공선성(multicollinearity) 검토를 실시하고, 선정된 독립변수를 대상으로 종속변수와의 상관관계를 분석하여 회귀식을 도출하였다.

4. 도로공사비 예측모델 설계

4.1 데이터 수집

본 연구에서는 일반도로공사비 예측을 하는 인공신경망모델 설계 및 회귀분석을 실시하기 위하여, 조달청 나라장터 국가종합전자조달시스템을 이용, 2003년에서 2010년까지 국토해양부 소속의 5대 지방관리청에서 공고한 신규 도로건설 공사 입찰공고를 대상으로 개략견적모델설계를 위한 실적자료를 수집하였다.

그림 5는 공사입찰 설명서를 이용하여 수집된 데이터의 종류를 나타낸다. 이 중, 계약형태, 개찰일시, 공사유형, 도로의 등급, 공사 지역, 도로의 총 길이, 도로의 폭, 터널 연장, 교량 연장, 교차로, 공사기간, 수요기관은 도로의 공사비에 영향을 미치는 인자이며, 추정금액, 추정가격, 관급금액, 기초금액, 낙찰금액은 영향인자에 의해 종속적으로 변한다.

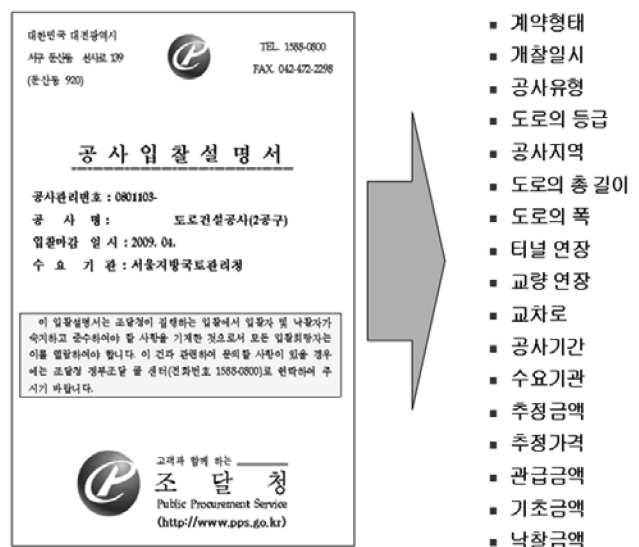


그림 5. 데이터 수집

본 연구는 일반 도로의 신설도로공사로 범위를 한정하여 데이터를 수집하였기 때문에 영향인자 중에 도로 유형, 도로의 등급은 입력값에서 제외하였다. 또한, 공사 금액을 결정하는데 가장 기본적인 값인 기초금액을 목표값으로 설정하였다.

모델의 입력값과 목표값을 설정 후, 모델을 구축하기 위해 수집한 자료에 대하여 수치화하였다(표 3). 먼저 정성적인 데이터인 계약방법, 개찰일시, 공사지역은 특징에 따라 0부터 정수로 설정하였으며, 정량적인 데이터인 도로 총 연장, 도로의 폭, 터널 연장, 교량 연장, 교차로 수, 공사기간은 고유 수치를 이용하였다. 또한, 출력값인 기초금액은 공사의 규모를 고려하여 억 단위로 설정하였다. 이 때, 기초금액을 인공신경망에 적용한 경우는 입력값에 개찰일시가 있는 점을 감안하여 공사비 지수를 이용하여 금액을 환산하는 것은 실시하지 않았으며, 회귀분석에서는 기초금액을 공사비 지수로 환산하여 사용하였다. 변환과정에서 표 3과 같이 변환하기 어려운 경우에 대해서는 샘플에서 제외시켰으며, 이러한 과정을 통해 2003년부터 2010년까지 총 93개의 샘플 데이터를 수집 및 변환하였다. 실적공사비가 도입된 2007년 이

전의 데이터 55개, 이후의 데이터 38개로 특정 개찰일시에 편중되지 않도록 샘플을 취득하였다.

표 3. 입력변수 및 출력변수 선택 및 값 설정

		영향인자	설정
입력변수	계약방법	1등급+PQ	0
		2등급+PQ	0
		3등급+PQ	0
		4등급+PQ	0
		시평액70%+PQ	0
		실적+PQ	0
		1등급	1
		2등급	1
		3등급	1
		4등급	1
		5등급	1
		시공능력공시역	1
		대안입찰	2
		일반경쟁	3
		개찰일시	년
	월		월
	공사지역	경상도	0
		전라도	1
		충청도	2
		경기도	3
		강원도	4
		제주도	5
		특별/광역시	6
		도로 총 연장	km
		도로의 폭	m
		터널 연장	km
		교량 연장	km
	교차로의 수	개	
	공사기간	월	
출력변수	기초금액	억	

4.2 인공신경망모델 설계

입/출력변수를 결정하여 그림 6과 같이 도로공사비 예측 인공신경망모델의 기본적인 구조를 설계하였다. 이 때, 인공신경망모델의 성능을 높이기 위해서 모델의 성능을 좌우하는 은닉층의 개수, 각 은닉층의 뉴런의 개수, 학습률, 모멘텀, 학습방법, 활성화 함수 등 여러 변수들을 적용할 문제에 맞게 적절하게 설정해주어야 한다. 하지만 이러한 변수의 변화에 따라 인공신경망모델의 성향이 어떻게 달라지는지 예측하기가 힘들기 때문에, 기본적으로 신경망 모델 설계시 결과값과 목표값의 오차를 최소화하기 위해 신경망의 각 변수를 조정해가며 결과값을 산출하여 목표값과 비교하는 시행착오과정을 거쳐 모델을 최적화 한다(Hegazy 등, 1994).

하지만, 시행착오방법은 여러 종류의 변수에 대하여 조건별로 모두 판단해야 하기 때문에, 많은 시간과 노력을 필요로 하고, 이러한 단점을 보완하기 위하여 최적 변수를 찾아

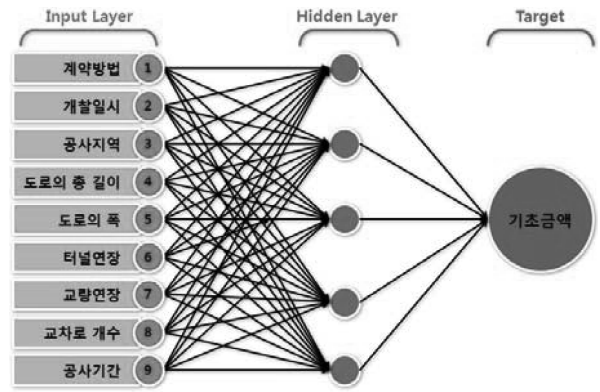


그림 6. 인공 신경망의 입력 및 목표값 설정

가는 과정을 유전알고리즘 등의 최적화 알고리즘을 사용하여 최적화된 변수들을 설정하는 연구들이 과거에 선행되었다(Hegazy와 Ayed, 1998; 김광희와 강경인, 2003; 김광희와 안성훈, 2007; 김광희 등, 2006).

본 논문에서는 최적 모델을 찾아내는 과정에 있어서 별도의 알고리즘을 사용하지 않고, 시행착오방법을 사용하여 최적 변수를 찾아 결과를 도출하였고, 시행착오방법을 이용하기 위한 변수 범위 및 함수를 표 4와 같이 설정하였다. 모델은 오류 역전파 알고리즘을 사용한 전방향 오류 역전파 신경망으로 지정하였으며, 은닉층의 개수는 1개, 은닉층과 출력층의 활성화 함수는 학습방법은 전방향 오류 역전파 신경망 각각 Tan-sigmoid/linear transfer function을 사용하였으며 모델의 성능평가는 평균제곱오차(mean squared error, MSE) 값을 이용하였다. 또한 샘플은 훈련과정, 입증과정, 시험과정에 각각 70%, 15%, 15%씩 분배하여 사용하였다.

표 4. 개략 공사비 예측을 위한 모델 설계

구분	설정
모델	전방향 오류 역전파 신경망 (feed-forward backpropagation network)
은닉층 개수	1
은닉층 뉴런 개수	1-20
은닉층 활성화 함수	Tan-sigmoid transfer function
출력층 활성화 함수	linear transfer function
성능평가	mean squared error (MSE)
훈련방법	Levenberg-Marquardt backpropagation
학습률	0.1-0.9
모멘텀	0.1-0.9
샘플선정	훈련 70%, 입증 15%, 시험 15%

이 때, 시행착오방법을 사용하기 위하여 조정할 변수는 은닉층의 뉴런의 개수, 학습률, 모멘텀으로 선정하여 최적화를 위한 범위를 정하였으며(김광희와 강경인, 2003; 김청용과 손재호, 2006), 은닉층 뉴런의 개수는 1개에서 20개까지 1개 단위로, 학습률과 모멘텀은 0.1에서 0.9까지 0.1 간격으로 범위를 설정하였다.

설계한 모델에 대하여 실험을 하기 위하여 The MATLAB 소프트웨어 V.7.8 버전의 Neural Network Toolbox™6을 이용하였다. Matlab의 Neural Network Toolbox는 Matlab

소프트웨어 내에서 제공하는 인공지능망 구현 틀로서 사용자가 사용하기 쉽도록 인터페이스를 구현해 놓았으며, 시행착오방법으로 변수들을 조정해가며 반복적으로 실험을 실시할 경우, Matlab의 기본함수들과 연동이 되어있기 때문에, 코딩작업으로 변수들을 변화시킨 모든 경우의 수에 대해 실험을 손쉽게 실시할 수 있고 엑셀과 연동하여 손쉽게 결과값을 이용할 수 있다.

4.3 다중회귀분석 모델 설계

본 연구에서는 도로공사비 예측 모델을 구축하기위해 2003년에서 2010년까지의 80건의 도로건설공사 입찰공고 자료를 대상으로 다중회귀분석을 수행하였다. 물가변동률 등을 고려하여 수집된 자료의 시점을 통일하기위해 건설공사비지수(2005년 평균=100)를 적용하여 종속변수인 기초금액을 동일 시점의 값으로 환산 적용하였다.

본 연구에서는 다중회귀분석을 수행하기 위해 통계소프트웨어인 SPSS 12.0를 사용하였고, 표 5와 같이 도로 총길이, 도로폭, 교량길이, 터널연장, 교차로, 공사기간을 후보 독립변수로 선정하여 변수들 간 상관분석을 통해 공선성을 파악하였으며, 단계별 선택법을 다중회귀분석 모델로 선정하여 분석을 수행하였다. 단계별 선택법은 각 단계별로 독립변수의 설명력이 높은 순서대로 변수를 채택하여 분석을 수행하는 방법으로, 본 연구에서는 변수들의 진입과 제거를 통제하기 위해 변수들 중 0.05보다 유의적인 변수는 포함되고, 새로운 변수의 진입에 따라 유의성이 0.10보다 떨어질 경우

표 5. 다중회귀분석 모델 변수 정의

구분		변수
독립변수	도로 총길이	x_1
	도로 폭	x_2
	교량길이	x_3
	터널연장	x_4
	교차로	x_5
공사기간	x_6	
종속변수	기초금액	y

표 6. 후보 독립변수 간 상관관계 분석

		도로 총길이	도로 폭	터널연장	교량길이	공사기간	교차로
도로 총길이	Pearson 상관계수	1	.092	.092	.173	.299(**)	.157
	유의확률		.209	.209	.063	.004	.082
도로 폭	Pearson 상관계수	.092	1	-.201(*)	-.144	.154	.081
	유의확률	.209		.037	.102	.086	.238
터널연장	Pearson 상관계수	.092	-.201(*)	1	.373(**)	.048	-.197(*)
	유의확률	.209	.037		.000	.336	.040
교량길이	Pearson 상관계수	.173	-.144	.373(**)	1	.222(*)	-.153
	유의확률	.063	.102	.000		.024	.088
공사기간	Pearson 상관계수	.299(**)	.154	.048	.222(*)	1	.076
	유의확률	.004	.086	.336	.024		.250
교차로	Pearson 상관계수	.157	.081	-.197(*)	-.153	.076	1
	유의확률	.082	.238	.040	.088	.250	

*상관계수는 0.05 수준에서 유의함.

**상관계수는 0.01 수준에서 유의함.

변수를 제거하도록 설정하였다(이학식과 임지훈, 2005).

표 6은 변수들 간 상관관계를 분석하기위해 Pearson 상관계수 및 유의확률을 도출한 결과이다. 일반적으로, 유의수준 0.05를 기준으로 0.9 이상의 상관계수를 보이는 독립변수들 사이에는 강한 상관관계로 인해 공선성 문제가 발생할 수 있다고 판단할 수 있으며(권호석 등, 2008), 분석 결과 독립변수 간 상관관계를 보이지 않는 것으로 분석되어 모든 후보 독립변수를 모두 선정하여 분석을 수행하였다.

5. 실험 및 결과

5.1 인공지능망 실험

설계한 인공지능망 실험에 앞서, 모델의 최적화를 위하여 입력변수의 영향도를 분석하였다. 기본적으로 설정된 10개의 입력변수(계약방법, 개찰년도, 개찰월, 공사지역, 도로 총 연장, 도로폭, 터널연장, 교량연장, 교차로) 중, 오차를 최소화 하는 입력변수의 종류를 판단하기 위하여, 10개 입력변수 중에서 1개를 제외한 후 모델의 성능평가를 실시하여 MSE 값을 비교하였다. 이 때, 입력변수 중 개찰일시에 관한 변수는 총 2가지가 있기 때문에, 개찰년도만 제거하였을 경우, 개찰월을 제거하였을 경우, 2가지 모두 제거하였을 경우로 나누어 실험을 실시하였다. 실험할 때의 조건은 표 4과 같으며, 모멘텀과 학습률은 각각 0.9, 0.1로 지정하여 각 조건당 50회씩 실험을 실시하였다.

실험결과(그림 7)를 보면, 입력변수를 모두 포함하였을 때에 비해서 오차가 크게 감소하는 경우는 없는 것으로 확인되었다. 또한, 입력변수 중 정량적인 성격을 가지고 있는 도로총연장, 도로폭, 터널연장, 교량연장에 대해서는 포함을 하지 않았을 경우 MSE가 급격하게 상승하였고, 특히 교량총연장을 입력변수에 포함시키지 않았을 경우, MSE가 크게 증가하였다. 입력변수를 조정하여 실시한 실험을 통하여, 사전 설정했던 입력변수를 모두 포함하여 모델을 설계하는 것으로 결론을 내렸고, 특히 도로 총연장, 도로 폭, 터널연장, 교량 연장은 꼭 포함해야 하는 요소라는 것을 확인하였다.

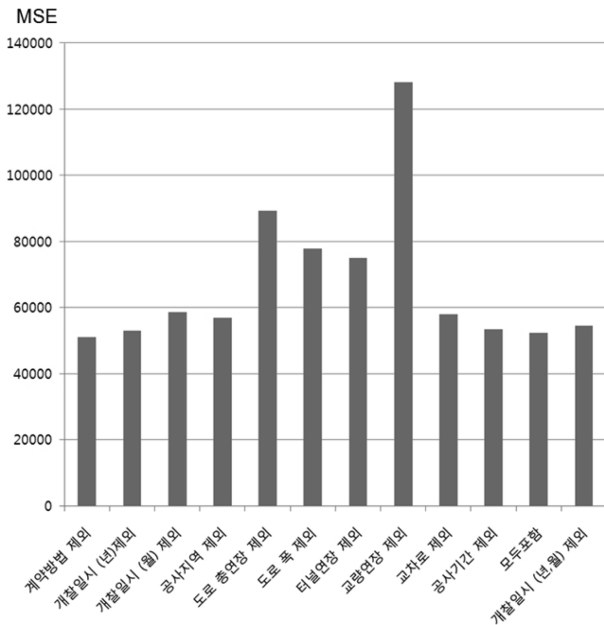


그림 7. 입력변수별 실험 결과값과 목표값의 평균 MSE

5.2 인공신경망모델 실험 결과

입력변수 설정 후, 표 4에서 설정한 은닉층 뉴런의 개수, 모멘텀, 학습률의 범위에 대하여 시행착오방식으로 실험을 실시하였다. 설정한 범위에 의하면 실험할 경우의 수는 은닉층 뉴런 개수의 경우 20개, 학습률과 모멘텀의 경우 각 9개로 하여 총 1620개의 경우가 발생한다. 1620개의 경우에 대하여 표 4와 같은 조건으로 인공신경망모델 실험을 실시

표 7. 성능 기준별 모델 변수 및 성능

성능 기준	MSE
은닉층 뉴런 개수 (개)	9
학습률	0.2
모멘텀	0.4
MSE (억 ²)	4435.3
평균 절대오차율 (%)	8.79

표 8. 시험 샘플에 대한 공사비 예측 결과

번호	목표값 (억)	결과값 (억)	오차 (억)	절대 오차율 (%)
1	548.3	619.5	-71.2	12.99
2	401.1	421.0	-19.9	4.97
3	1241.9	1178.6	63.3	5.10
4	686.7	791.2	-104.5	15.21
5	340.8	293.8	47.0	13.78
6	1147.4	1063.0	84.4	7.36
7	1128.7	1107.0	21.7	1.92
8	947.9	984.0	-36.1	3.81
9	922.8	1000.5	-77.7	8.42
10	1018.6	914.8	103.8	10.19
11	1240.6	1243.5	-2.9	0.23
12	343.0	429.4	-86.4	25.18
13	845.6	802.2	43.4	5.13
평균	831.8	834.5	-2.7	8.79

하였으며 결과 은닉층 뉴런의 개수 9개, 학습률 0.2, 모멘텀 0.4에서 최소의 MSE를 가졌고, 평균절대오차는 8.79%로 계산되었다(표 7). 표 8은 표 7의 조건인 경우에 13개의 시험 샘플에 대한 공사비 예측의 오차율을 가리킨다.

5.3 다중회귀분석 모델 결과

본 연구에서는 회귀식의 설명력을 나타내는 결정계수 R²을 활용하여 단계별 선택법에 의해 도출된 회귀 모형의 유효성을 분석하였다. 표 9에서 볼 수 있듯이 모형 4가 0.851로 가장 큰 값을 보이며, 이는 종속변수의 정보 중 85.1%가 독립변수의 변동으로 설명될 수 있음을 나타낸다. 다음으로 회귀식의 유의성을 판단하기 위해 표 10과 같이 분산분석표를 분석하여 단계별 선택법에 의한 모형 4가지 모두 유의확률이 95% 신뢰구간에서 유의성을 확보할 수 있는 일반적인 유의수준인 0.05보다 작은 것으로 도출되어 통계적으로 유의함을 볼 수 있다.

표 9. 다중회귀모형 요약

모형	R	R ²	수정된 R ²	추정값의 표준오차
1	.711(a)	.505	.499	351.9384296
2	.822(b)	.676	.667	286.6683447
3	.875(c)	.766	.757	245.1701027
4	.923(d)	.851	.843	196.7969093

a예측값: (상수), 교량길이

b예측값: (상수), 교량길이, 도로 총길이

c예측값: (상수), 교량길이, 도로 총길이, 터널연장

d예측값: (상수), 교량길이, 도로 총길이, 터널연장, 도로 폭

표 10. 분산분석표

모형	제공합	자유도	평균제공	F	유의확률	
1	회귀모형	9.86×10 ⁶	1	9.86×10 ⁶	79.571	.000(a)
	잔차	9.66×10 ⁶	78	1.24×10 ⁵		
	합계	1.95×10 ⁷	79			
2	회귀모형	1.32×10 ⁷	2	6.59×10 ⁶	80.246	.000(b)
	잔차	6.33×10 ⁶	77	8.22×10 ⁴		
	합계	1.95×10 ⁷	79			
3	회귀모형	1.50×10 ⁷	3	4.98×10 ⁶	82.898	.000(c)
	잔차	4.57×10 ⁶	76	6.01×10 ⁴		
	합계	1.95×10 ⁷	79			
4	회귀모형	1.66×10 ⁷	4	4.15×10 ⁶	107.233	.000(d)
	잔차	2.90×10 ⁶	75	3.87×10 ⁴		
	합계	1.95×10 ⁷	79			

a예측값:(상수), 교량길이

b예측값:(상수), 교량길이, 도로 총길이

c예측값:(상수), 교량길이, 도로 총길이, 터널연장

d예측값:(상수), 교량길이, 도로 총길이, 터널연장, 도로 폭

4가지 모형의 결정계수와 분산분석표를 종합적으로 분석하여 유의수준 0.05이하를 만족시키며, 가장 큰 설명력을 보이는 모형 4를 공사비 예측 회귀식을 도출하였으며, 표 11의 계수 분석 결과를 바탕으로 식 (3)과 같이 공사비 예측 모델을 선정하였다. 또한, 선정된 모형의 독립변수들 간 다중

표 11. 다중회귀모형 계수

	비표준화계수		표준화계수	t	유의확률	공선성 통계량	
	B	표준오차	베타			공차	VIF
(상수)	-797.547	134.125		-5.946	.000		
교량길이	430.841	38.285	.549	11.253	.000	.834	1.199
도로 총길이	60.249	7.372	.373	8.172	.000	.954	1.048
터널연장	145.783	18.853	.377	7.733	.000	.836	1.196
도로 폭	42.018	6.411	.301	6.554	.000	.939	1.065

공선성을 판단하기 위해, 한 독립변수가 다른 독립변수들에 의해 설명되는 정도를 나타내는 지수인 공차한계와 분산팽창요인(variance inflation factor; VIF)을 분석하여 일반적 기준인 공차한계 0.10 이상, 분산팽창요인 10 이하(이학식과 임지훈, 2005)를 만족시켜, 공선성에 문제가 없음을 확인할 수 있다.

$$y = -797.547 + 430.841x_3 + 60.249x_1 + 145.783x_4 + 42.018x_2 \quad (3)$$

표 12는 산출된 회귀식을 검증하기 위해 13개의 시험 샘플자료와 회귀식에 의한 추정값을 비교한 결과이다.

표 12. 시험 샘플에 대한 공사비 예측 결과

번호	목표값 (억)	결과값 (억)	오차 (억)	절대오차율 (%)
1	413.5	503.5	90.0	21.77
2	302.5	383.4	80.9	26.74
3	1228.4	1071.4	-157.0	12.78
4	519.0	694.4	175.3	33.78
5	255.5	194.6	-60.8	23.81
6	1077.4	983.1	-74.3	8.75
7	1093.7	1057.2	-36.5	3.34
8	956.5	981.0	24.5	2.56
9	691.8	767.0	75.2	10.88
10	1020.6	926.6	-94.0	9.21
11	1364.8	1180.4	-184.4	13.51
12	254.8	528.0	273.2	107.20
13	848.1	805.8	-42.3	4.99
평균	771.3	775.1	3.8	21.49

5.4 결과 비교

인공신경망의 성능을 검증하기 위하여, 시험에 쓰인 13개의 샘플에 대하여 인공신경망모델의 결과를 다중회귀분석결과와 국토해양부의 기준을 적용한 결과와 비교를 실시하였다(표 13). 비교 결과, 인공신경망의 절대평균오차는 8.79%로, 다중회귀분석결과와 국토해양부기준을 적용한 결과인 21.49%, 80.32%에 비하여 상대적으로 성능이 우수함을 발견할 수 있었다.

표 13. 인공신경망, 회귀분석, 국토해양부기준의 결과 비교

	인공신경망	회귀분석	국토해양부기준
절대평균오차	8.79 %	21.49 %	80.32 %

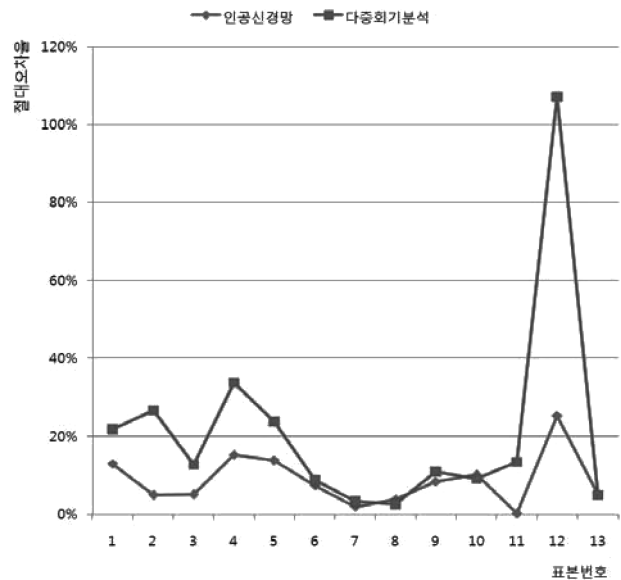


그림 8. 인공신경망모델과 회귀분석 결과의 샘플별 비교

인공신경망모델과 회귀분석을 적용한 결과를 그림 8과 같이 13개의 샘플별로 분석을 해보면, 인공신경망모델이 회귀분석을 적용한 결과에 비하여 전반적으로 작은 절대오차율을 가졌다. 또한, 회귀분석의 경우 107.20%의 절대오차율을 갖는 12번 샘플을 이상값으로 간주한 후 12개 샘플의 평균 절대오차율을 구한 결과도 14.34%로 인공신경망모델의 절대오차율이 여전히 더 작은 값을 가졌다. 이와 같은 결과로 인공신경망이 다중회귀분석, 도로공사비 현행 국토해양부의 기준에 비해 기획단계의 도로공사비 예측면에서 더 우수하다는 것을 알 수 있었다.

6. 결 론

본 연구에서는 현행 일반도로건설 개략 공사비 예측의 한계를 보완하기위해 인공신경망을 적용하여 공사 입찰을 위한 설계도면 작성 이후 물량산출서에 의해 산정된 상세 공사비 및 공사비를 결정하는 변수를 포함한 과거 실적 자료를 바탕으로 기획단계에서 공사비 예측이 가능한 도로 공사비를 개략전적 하는 모델을 제안하였다. 최소의 오차를 발생시키는 개략전적 공사비 예측 인공신경망모델을 제안하기 위해 시행착오방식을 적용하여 입력변수의 종류, 은닉층 뉴런의 개수, 학습률, 모멘텀을 결정하였으며, 결정된 인공신경망모델의 예측 결과를 다중회귀분석결과 및 현행 사용되는 국토해양부의 결과와 비교하여 정확도를 검증하였다. 방법별 절대오차의 비교결과 인공신경망모델이 공사비 예측에서 정

확도가 더 우수함을 확인할 수 있었고, 도로공사분야에 대한 배경지식 없이 입/출력변수를 조정함으로써 공사추진 단계별로 공사비 예측을 쉽게 적용할 수 있다는 점과 별도의 분류 없이 정성적 데이터와 정량적 데이터를 함께 사용할 수 있다는 점을 고려할 때, 기획단계의 도로공사비 견적 분야에 효율적으로 사용될 수 있을 것으로 예상된다. 다만 본 연구에서 적용한 모델 변수 최적화 방법인 시행착오방법은 모든 경우의 수에 대해 각각 실험을 실시해주어야 하기 때문에 상대적으로 시간과 노력이 많이 소요된다는 단점을 발견할 수 있었다. 향후에 김광희와 강경인(2003), Hegazy와 Ayed (1998) 등이 연구한 것과 같이 시행착오과정을 대신하여 유전알고리즘 등의 최적화 알고리즘을 적용하여 인공지능망의 변수를 효율적으로 최적화 할 수 있는 방법을 적용하는 연구가 진행된다면, 기획단계의 도로공사비 예측을 하는 인공지능망모델이 넓게 활용될 수 있을 것이다.

참고문헌

곽수남, 김두연, 김병일, 최석진, 한승헌(2009) 가용정보를 활용한 기획 및 설계초기 단계의 도로 공사비 예측모델. **한국건설관리학회 논문집**, 한국건설관리학회, 제10권, 제4호, pp. 87-100.

국토해양부(2009) **도로업무편람**. 국토해양부.

김광희, 강경인(2003) 공동주택 프로젝트의 초기 공사비 예측을 위한 신경망 학습에 유전자 알고리즘을 적용한 모델에 관한 연구. **대한건축학회 논문집(구조계)**, 대한건축학회, 제19권, 제10호, pp. 113-142

김광희, 안성훈, 조형근(2006) 신경망과 유전자알고리즘을 이용한 공사비예측 모델의 예측 정확도 비교에 관한 연구. **대한건축학회 논문집(구조계)**, 대한건축학회, 제22권, 제3호, pp. 111-118.

김광희, 안성훈(2007) 공사단계별 공사비 영향변수의 선택방법과 데이터 개수와의 상관관계에 관한 연구. **대한건축학회논문집(구조계)**, 대한건축학회, 제23권, 제4호, pp. 129-137.

김선국, 신진수, 구인완, 김용구(2000) 계획단계 도로공사의 통계

적 코스트 모델. **대한토목학회논문집**, 대한토목학회, 제20권 제2D호, pp. 171-180

김청용, 손재호(2006) 교육시설의 개념단계 공사비예측을 위한 인공지능망모델 개발에 관한 연구. **한국건설관리학회논문집**, 한국건설관리학회, 제7권, 제4호, pp. 91-99.

김해경, 윤진희(2004) **확률과 통계-이론적 기초**. 경문사.

박홍태(2006) **건설공사를 위한 계약 및 적산**. 구미서관.

윤유중, 우성권(2006) 회귀분석을 이용한 고속도로 건설공사 예비비의 산정. **대한토목학회논문집**, 대한토목학회, 제26권 제4D호, pp. 617-626.

이학식, 임지훈(2005) **SPSS 12.0 매뉴얼**. 법문사.

한학용(2009) **패턴인식 개론**, 한빛미디어.

Bode, J. (1998) Neural networks for cost estimation. *Cost Engineering*, AACE, Vol. 40, No. 1, pp. 25-30.

Hegazy, T., and Ayed, A. (1998) Neural network model for parametric cost estimation of highway projects. *Journal of Construction Engineering and Management*, ASCE, Vol. 124, No. 3, pp. 210-218.

Hegazy, T., Moselhi, O. and Fazio, P. (1994) Developing practical neural network applications using back-propagation. *Journal of Microcomputers in Civil Engineering*, Blackwell Publisher, Vol. 9, No. 2, pp. 145-159.

Oberlender, G.D. and Trost, S.M. (2001) Predicting accuracy of early cost estimates based on estimate quality. *Journal of Construction Engineering and Management*, ASCE, Vol. 127, No. 3, pp. 173-182.

Petrooutsatou, C., Lambropoulos, S., Pantouvakis, J.P. (2006) Road tunnel early cost estimates using multiple regression analysis. *Operational Research*, Springer Berlin, Vol. 6, No. 3, pp. 311-322.

Wilmot, C.G. and Cheng, G. (2003) Estimating future highway construction costs. *Journal of Construction Engineering and Management*, ASCE, Vol. 129, No. 3, pp. 272-279.

Wilmot, C.G. and Mei, B. (2005) Neural Network Modeling of Highway Construction Costs. *Journal of Construction Engineering and Management*, ASCE, Vol. 131, No. 7, pp. 765-771.

(접수일: 2011.8.2/심사일: 2011.9.8/심사완료일: 2011.9.8)