

한국과 미국간 항공기 탑승객 수 예측을 위한 뉴럴네트워의 응용[†]

남경두

건국대학교 상경대학 무역학과

ABSTRACT

In recent years, neural networks have been developed as an alternative to traditional statistical techniques. In this study, a neural network model was compared to traditional forecasting models in terms of their capabilities to forecast passenger traffic for flights between U.S. and Korea. The results show that the forecasting ability of the neural networks was superior to the traditional models. In terms of accuracy, the performance of the neural networks was quite encouraging. Using mean absolute deviation, the neural network performed best. The new technique is easy to learn and apply with commercial neural network software. Therefore, airline decision makers should benefit from using neural networks in forecasting passenger loads.

1. 서론

항공회사에서 탑승객 수의 정확한 예측은 성공적인 회사경영을 위하여 필요하다. 항공회사의 의사결정자들은 좌석배정, 여름 성수기동안의 추가적인 인력확보, 광고예산편성 등을 위하여 정확한 탑승객 수의 예측을 필요로 한다. 또한 각 공항 관계자들은 공항 터미널 시설 확충과 같은 의사결정을 위하여 탑승객 수의 정확한 예측을 사용한다[8].

탑승객 수의 예측을 위한 시계열 통계분석방법들은 회귀분석(regression analysis)과 지수평활법(exponential smoothing)등 다양하다. 그러나 이들 통계분석방법들은 다음과 같은 단점을 가지고 있다.

첫째, 이들 통계분석방법들은 몇 가지 전제가 되는 가정들이 있다. 예를 들면 탑승객 수 예측을 위하여 가장 널리 쓰이는 통계분석방법들중에 하나는 회귀분석이다[8]. 회귀분석의 시행은 다음과 같은 회귀모형의 측정오차에 대한 전제조건이 충족되어야 한다.

- (1) 측정오차들은 서로 독립이 되어야 한다.
- (2) 측정오차들은 정규분포를 가쳐야 한다.
- (3) 측정오차들의 분산은 일정해야 한다.

[†] 이 논문은 건국대학교 신입교원 연구정착금 지원에 의하여 이루어졌음

이들 전제조건들의 약간의 위반은 회귀분석의 예측능력에 커다란 영향은 미치지 않으나 이들 가정들이 크게 위반되었을 때에는 회귀분석모델의 수행에 커다란 영향을 미친다 [4] [10]. 역설적으로 그런 가정들의 위반은 회귀모형의 시행 후에만 발견될 수 있다. 이런 이유 때문에 분배 가정을 덜 요구하는 비모수통계기법과 같은 대체 기법의 개발이 고려되고 있다.

둘째, 실제 탑승객 수 예측을 위한 시계열 자료는 여러 가지 복잡한 패턴을 포함하고 있기에 기존 통계기법을 통한 적절한 통계모델을 만드는 절차가 복잡하다 [1]. 셋째, 기존의 통계분석에서는 주어진 문제에 대해서 여러 가지 통계적 모델이 존재한다. 기존의 여러 가지 통계기법들 중에서 주어진 자료에 대한 가장 좋은 결과를 제공하는 통계모델의 선택을 위한 지침이 마련되어 있지 않다. 그래서 항공회사의 의사결정자들은 주관적인 경험에 의해서 각종 통계적 기법과 모델을 선택하고 있다.

최근에 컴퓨터의 발달에 따라 컴퓨터를 응용으로 한 여러 가지 예측을 위한 통계분석방법이 개발되고 있다. 전통적인 통계분석의 약점을 보완한 새로운 컴퓨터-인센티브(computer-incentive) 통계기법들 중에 하나가 뉴럴네트웍이다. 뉴럴네트웍은 인공지능(artificial intelligence)테크닉의 한 분야로서 인간의 두뇌기능을 모방한 수학적 모델을 근거로 한 컴퓨터 시스템이다[12]. Wasserman[17]에 의하면 뉴럴네트웍은 인간처럼 경험으로부터 배울 수 있고 과거의 패턴으로부터 새로운 일반화된 패턴을 이끌어 낼 수 있으며 결측(missing) 또는 잡음(noisy)자료로부터 어떤 근본적인 특성을 도출할 수도 있다. 추가적으로 뉴럴네트웍은 전통적인 통계기법들처럼 분배에 대하여 같은 가정을 요구하지 않는다. 뉴럴네트웍은 분배 모양에 대하여 상대적으로 약한 가정을 요구하기 때문에 전통적인 통계기법의 대안으로 사용되고 있다.

이에 본 연구에서는 항공기 산업에서 효율적인 의사결정을 위한 뉴럴네트웍을 소개하고 응용분야로서 한국과 미국간의 국제항공기 탑승객 수의 예측을 위한 기존 시계열 통계분석방법(회귀분석과 지수평활법)과 뉴럴네트웍의 수행을 비교 분석하는데 그 목적이 있다. 또한 이와 같은 연구 결과를 통하여 관련분야의 의사결정자에게 새로운 분석도구를 제공하고 동시에 이 분야에 관심을 가지고 연구의 뜻을 두고 있는 관계자들에 대해 이론적이고 실제적인 연구자료로 활용케 하는 데에 이차적인 목적이 있다. 만약 뉴럴네트웍의 실용가치가 본 연구를 통해서 입증되던 항공산업에서의 의사결정자들은 더욱더 정확한 예측을 위하여 이 새로운 컴퓨터 테크닉인 뉴럴네트웍의 사용을 적극적으로 고려해야 한다. 항공기 탑승객 수의 예측을 위한 뉴럴네트웍의 응용은 뉴럴네트웍의 이론적 배경의 이해를 요구하지 않았다. 이 새로운 컴퓨터 테크닉은 기존의 통계기법에 비해서 배우기가 쉽고, 현재 미국에서 시판중인 뉴럴네트웍의 소프트웨어에 의해서 쉽게 사용할 수 있다.

2. 뉴럴네트웍 (Neural Network)

최근에 인공지능의 탐구의 한 결과로서 뉴럴네트웍이 개발되었다. 인공지능 연구자들은 인간의 신경계와 유사한 독립적으로 생각할 수 있고 여러 가지 현실문제를 해결할 수 있는 컴퓨터모델을 개발하였다. 이 인공지능 개발의 한 결과가 뉴럴네트웍이다. 뉴럴네트웍은 인간의 두뇌와 유사한 기능을 하는 컴퓨터 모델이다.

뉴럴네트웍은 연결 가중치(weighted connection)를 사용하여 서로 상호 작용하는 뉴런(neurons) 또는 노드(nodes)라고 불리는 단순 처리 단위로 구성된 다이내믹 시스템이다. 일반적으로 뉴럴네트웍은 입력층(input layer), 출력층(output layer), 하나 이상의 은닉층(hidden layer)의 계층구조를 갖는다. 비록 은닉층의 수에 대한 이론적인 제한은 없으나 대부분의 문제를 푸는데는 한 개의 은닉층이면 충분하다. 입력층에 있는 입력값에 가중치가 정해지고 은닉층으로 전달된다. 은닉층에 있는 뉴런들은 입력층에서 전달된 각각의 입력값에 각각의 가중치를 곱한 합계를 기준으로 은닉값을 계산한다. 같은 방법으로 은닉층에 있는 계산된 은닉값에 가중치가 정해지고 최종 결과를 얻기 위하여 출력층으로 전달된다. 뉴럴네트웍은 서로 연결된 가중치(weight)를 조정함으로써 학습된다. 뉴럴네트웍에 의해서 생성된 출력값과 실제값(desired output) 사이의 오차가 계산되고 정해진 최소한도의 오차를 얻을 때까지 연결된 가중치를 수정해 나간다.

2.1 인공뉴런

뉴럴네트웍에서 계산절차는 인공뉴런의 활동과정을 조사함으로써 쉽게 이해될 수 있다. x_1, x_2, \dots, x_p 를 p 개의 입력패턴이라 하고 w_1, w_2, \dots, w_p 를 각각의 입력패턴과 연결된 가중치라고 하자. 수학적으로 p 개의 입력패턴과 가중치는 (x_1, x_2, \dots, x_p) 와 (w_1, w_2, \dots, w_p) 처럼 벡터로서 표현할 수 있다. 입력 신호의 합계(NET)는 두개의 벡터의 내적(inner product)이다 [12]. 그래서 NET는 다음과 같이 정의된다.

$$NET = \sum w_i x_i \quad (1)$$

다음 단계로 NET는 뉴런의 출력 신호, $f(NET)$ 를 생성하기 위하여 활성화함수(activation function) f 에 의해 변형된다. 지금까지 이용할 수 있는 활성화함수는 S자 형태의 시그모이드(sigmoid)함수, 쌍곡선의 탄젠트(hyperbolic tangent), 계단함수(hard limiter)등 다양하다[12]. 이들 중 시그모이드함수가 가장 널리 쓰이고 있다. 시그모이드 함수는 출력값의 범위를 0과 1사이로 제한하고 있고 [18], NET의 함수로써 다음과 같이 나타난다.

$$f(NET) = 1/(1+e^{-NET}) \quad (2)$$

2.2 뉴럴네트웍의 훈련

뉴럴네트웍의 시행에는 두개의 주요 과정이 있다. 하나는 훈련(training)과정이고 다른 하나는 검증(testing)과정이다. 훈련과정은 뉴럴네트웍이 생성하는 오차(error)에 대하여 연결 가중치를 수정하는 절차이다. 검증과정은 뉴럴네트웍이 입력된 패턴을 기억하기보다는 그 패턴에 대한 일반화를 도출할 수 있는가에 대하여 검증하는 절차이다. 훈련의 두 가지 유형으로는 지도훈련과정(supervised training)과 자율훈련과정(unsupervised training)이 있다. 지도훈련과정에서는 뉴럴네트웍을 통해 나온 출력값이 기대값과 비교가 되며 이들 두 값의 차를 오차라고 부른다. 이 오차는 뉴럴네트웍을 통하여 역전과되며 가중치는 그 오차를 줄이기 위하여 수정된다[15]. 지도훈련과정에서는 기대값이 필요치 않는다. 그래서 뉴럴네트웍의 출력값은 기대값과 비교되지 않는다. 뉴럴네트웍을 훈련시키기 위하여 여러 가지 알고리즘이 개발되고 있다[30]. 이 중 가장 흔히 쓰이

는 알고리즘이 백프로퍼게이션(backpropagation)이다. 백프로퍼게이션 알고리즘은 Werbos[18], Parker[13], Rumelhart, Hinton, and Williams[14] [15] 등에 의해 여러 차례 독립적으로 개발되었다. 백프로퍼게이션은 네트워크를 통하여 출력오차를 역으로 전파하고 각층에서 가중치를 수정함으로써 은닉층을 훈련시킨다. 백프로퍼게이션은 다층 뉴럴네트워크를 훈련시키기 위한 체계적 방법이고 뉴럴네트워크를 쉽게 응용분야로 이용될 수 있는 수단을 제공한다. 백프로퍼게이션은 일반화 델타 규칙(generalized delta rule)을 사용한다. 일반화 델타 규칙 p 번째의 입력/출력의 쌍(pair)이 제시되는 경우에 노드 i 에서 노드 j 로의 연결가중치의 변화는 다음과 같이 표현된다.

$$\Delta_p w_{ji} = \eta (t_{pj} - o_{pj}) X_{pi} = \eta \varepsilon_{pj} X_{pi} \quad (3)$$

여기서 η 은 학습속도(learning rate)을 결정하는 상수, t_{pj} 는 p 번째 출력패턴의 j 번째 성분을 위한 목표 출력, o_{pj} 는 p 번째 입력패턴으로부터 뉴럴네트워크가 계산한 j 번째 성분을 위한 실제 출력, X_{pi} 는 p 번째 입력패턴의 i 번째 성분의 값, $\varepsilon_{pj} = t_{pj} - o_{pj}$ 는 목표출력과 실제출력의 오차, $\Delta_p w_{ji}$ 는 패턴 p 의 입력층 i 번째 성분에서 j 번째 성분으로의 연결가중치의 변화량이다. 진동없이 학습속도를 증가시킬 수 있는 수정된 일반화 델타 규칙은 다음과 같다.

$$\Delta w_{ji}(t+1) = \eta \varepsilon_{pj} O_{pi} + \alpha \Delta w_{ji}(t) \quad (4)$$

여기서 α 는 모멘텀(momentum)을 결정하는 상수이다. 일반화델타규칙의 더욱 상세한 수학적 묘사는 Rumelhart, Hinton, Williams[15]의 연구논문을 참조할 수 있다.

백프로퍼게이션 알고리즘의 단점중의 하나는 사용자가 뉴럴네트워크를 훈련시키기 위한 충분한 전문지식이 필요하다는 것이다. 이 전문지식이란 사용자가 어떤 주어진 문제에 대하여 은닉층의 수, 은닉층 뉴런의 수, 학습속도, 그리고 모멘텀을 결정해야 하는 것이다. 이들 모수(parameter)들의 선택에 대해서 현재까지는 구체적인 규칙 또는 절차가 정해져 있지 않다. 은닉층은 뉴럴네트워크의 입력층과 출력층 사이에 있는 중간층으로 외부 환경과 직접적으로 연결되어 있지 않다[15]. 입력과 출력 패턴들이 비선형이거나 복잡한 구조를 가지고 있거나 선형으로 분리될 수 없는 자료들도 은닉층에 의해서 쉽게 분리될 수 있다. 대부분 백프로퍼게이션 네트워크에서 은닉층의 뉴런의 수는 보통 입력과 출력 뉴런의 수의 총 합계 또는 그 보다 적은 수로 결정된다. 은닉층에서 뉴런의 수가 너무 적으면 네트워크를 훈련시키는데 요구되는 학습 반복 횟수를 증가시키고 뉴럴네트워크의 예측능력을 감소시킨다. 한편 은닉층에서의 너무 많은 뉴런의 사용은 훈련시간을 연장하고 뉴럴네트워크가 훈련데이터(training data)를 일반화하기보다는 오히려 기억하게 할 수도 있다[9]. 일반적으로 훈련데이터가 잡음이 많고 명확한 패턴을 가지고 있지 않으면 많은 수의 은닉 뉴런이 필요로 한다. 하지만 훈련자료의 패턴이 명확하면 적은 수의 은닉 뉴런이라도 뉴럴네트워크를 충분히 훈련시킬 수 있다.

학습속도(learning rate)는 0과 1사이의 범위를 가지고 있으며 현재의 가중치 벡터에서 목표의 벡터로 가기 위한 수렴속도를 조절하는 수단이다. 최급하강법(gradient descent method)은 학습속도가 적으면 천천히 기울기를 따라가지만 학습속도가 너무 크면 종종 목표의 가중치 벡터로 가지 않고 목표없이 진동한다. 진동없이 학습을 보다 빠르게 하는 방법은 모멘텀을 포함하는 것이다. 모멘텀은 가중치 공간에서 현재 하강하는 방향으로 과거 가중치를 수정해 나간다. 즉 연결 가중치를 수정하는데 있어서 목표점을 찾지 못하고 진동하기보다는 기울기를 따라가는 방향으로 가중치를 수정한다. 따라서 모멘텀항(momentum term)의 적당한 사용은 훈련시간을 단축시키고 훈련

의 안정성을 향상시켜 학습속도를 진동없이 효과적으로 증가시킬 수 있다. 학습속도처럼 모델은 0과 1사이의 값을 선택할 수 있다.

2.3 항공산업에서 뉴럴네트워크의 사용

뉴럴네트워크의 응용은 아주 다양하다. 예를 들면 Dutta와 Shekhar는 회사의 채권분류를 예측하는 분류문제로 뉴럴네트워크를 응용하였다[6]. Duliba는 운송산업에서 수행을 정확하게 예측하기 위한 수단으로 회귀분석과 뉴럴네트워크의 수행을 비교하였다[5]. Chang, Rapiraju, Whiteside, Hwang은 강우량을 예측하기 위하여 뉴럴네트워크를 시계열 분석모델과 비교하였다[3]. 현재까지 알려진 뉴럴네트워크의 응용은 패턴 분류와 예측을 위한 통계적 문제들을 포함한다. 탑승객 수 예측을 위한 뉴럴네트워크의 응용은 전통적인 통계기법의 여러 가지 문제점에 기인한다. 전통적인 통계기법에 비하여 뉴럴네트워크는 다음과 같은 경우에 장점을 지닌다[2].

- *데이터의 분배가 알려져 있지 않거나 또는 비선형을 나타낼 경우
- *모델 가정이 위반될 경우
- *이상값(outliers) 또는 부적절한 데이터가 존재하는 경우
- *입력변수가 많은 경우
- *왜곡된 데이터가 분석을 어렵게 하는 경우
- *선택된 모델이 부적절한 경우
- *온라인 의사결정이 필요한 경우

또한 뉴럴네트워크는 그 모델의 구조적 형태(structural form)를 사전에 결정할 필요가 없다. 예를 들면, 통계적 예측분석에서 어떤 주어진 문제에 대하여 선택할 수 있는 통계적 모델이 너무 다양하다. 뉴럴네트워크에서는 훈련을 위하여 입력과 출력 데이터만 필요하다. 더구나 뉴럴네트워크는 전통적인 통계기법 사용시 부적절한 여러 가지 통계적 예측문제에 적절히 사용될 수 있다. 예를 들면, 항공산업분야에서의 뉴럴네트워크의 사용은 탑승객 수 예측, 좌석배정, 예약율 예측 등을 비롯하여 통계적 분석과 관련된 여러 분야에 응용할 수 있다.

뉴럴네트워크와 회귀분석과의 유사점과 상이점은 다음과 같다. 유사점으로는 뉴럴네트워크와 회귀분석은 주어진 입력 변수의 값을 가지고 출력변수의 값을 예측하는 것이다. 추가적으로 뉴럴네트워크의 입력층 뉴런은 회귀분석에서 독립변수에 상응하고 출력층 뉴런은 종속변수에 상응한다. 그러나, 뉴럴네트워크는 그 자체의 독특한 특징을 가지고 있다. 첫째, 뉴럴네트워크는 동시에 한개 이상의 종속변수를 사용할 수 있다. 그러나 회귀분석은 오직 한 개의 종속변수만 사용할 수 있다. 둘째, 뉴럴네트워크는 최급하강법을 통하여 암암리에 자동적으로 데이터 변환(transformation)을 수행할 수 있는 반면, 회귀분석은 사용자가 변환을 명시하여야 하거나 또는 다른 대체되는 모델을 제시하여야 한다[11]. 셋째, 뉴럴네트워크는 오차를 줄이기 위하여 연결 가중치를 지속적으로 적응(adaptive)하는 반면 회귀분석은 모든 훈련데이터를 적응(adaption)없이 동시적으로 처리한다[5]. 마지막으로 전통적인 회귀분석에서는 독립변수가 유의한 설명변수인지 확인하기 위하여 t-통계량의 값을 평가한다. 유사한 방법으로 뉴럴네트워크에서는 입력 노드(node)로부터 계산되어진 모든 연결 가중치를 합산함으로써 해당하는 입력 변수의 중요성을 평가할 수 있다.

뉴럴네트웍은 여러 가지 단점을 가지고 있다. 첫째, 뉴럴네트웍은 훈련을 위하여 일정량 이상의 데이터가 필요하다. 특히 주어진 훈련패턴에 잡음이 존재하는 경우에는 상대적으로 많은 양의 데이터가 필요하다. 그러나 이런 경우에는 회귀분석을 포함하여 대다수의 예측모델에도 적용되는 같은 단점이다. 둘째, 모든 뉴럴네트웍은 학습절차가 요구된다[2]. 셋째, 훈련된 뉴럴네트웍의 내부구조는 뉴럴네트웍을 통해 나온 출력값이 어떻게 얻어졌는가에 대한 구체적인 단계를 추적하기가 어렵다[7].

3. 방법론

3.1 데이터 수집

항공기 탑승객 수의 예측을 위한 뉴럴네트웍과 전통적인 시계열 통계기법과의 수행비교분석을 위하여 1982년 1월부터 1993년 12월 동안 한국과 미국간 월별 탑승객 수를 위한 데이터가 미국에 있는 Volpe National Transportation System Center[16]로부터 수집되었다. 월별 탑승객 수를 위한 데이터의 시도표(time plot)는 월별 계절성(seasonality)을 가지고 위로 향하는 추세(trend)를 보여준다. 이 데이터는 다시 훈련과정과 검증과정을 위하여 두개의 데이터 세트(set)로 양분되었다. 훈련과정을 위한 훈련데이터 세트는 회귀분석과 지수평활모델을 생성하고 훈련과정동안 뉴럴네트웍에 여러가지 패턴을 제공하기 위하여 사용되었다. 검증과정을 위한 검증데이터 세트는 회귀분석, 지수평활법, 뉴럴네트웍의 예측능력을 검증하기 위하여 사용되었다. 표본크기가 다른 3년, 6년, 9년의 세 가지 훈련데이터세트가 선택되었다. 예를 들면 표1에서 제시된 것처럼 1991년도를 예측하기 위해서 1988년1월-1990년 12월, 1985년 1월-1990년 12월, 1982년 1월-1990년 12월의 세 가지의 다른 기간이 훈련데이터로 사용되었다.

정확도(accuracy)는 계산된 값이나 추정값이 기대값에 얼마나 가까이 있는가 하는 측도를 말한다. 즉, 정확도는 예측된 값과 실제값사이의 차의 측도이다. 본 연구에서의 정확도는 뉴럴네트웍, 회귀분석, 지수평활모델의 수행비교를 위하여 평균절대편차(MAD: mean absolute deviations)에 의해 측정된다. 평균절대편차는 각 잔차(residual)의 절댓값들의 평균값이다. 또한 본 연구에서는 각예측모델로부터 얻어진 각평균절대편차의 차분(difference)에 대한 통계적 증명을 얻기 위하여 F-검증을 사용한다.

<표 1> 훈련데이터와 검증 데이터 세트

훈련데이터			검증데이터
3년	6년	9년	
1988-1990	1985-1990	1982-1990	1991
1989-1991	1986-1991	1983-1991	1992
1990-1992	1987-1992	1984-1992	1993

3.2 가변수(dummy variable)를 가지는 시계열 회귀분석

본 연구에 사용된 데이터는 계절성을 가지고 지속적으로 위로 향하는 선형 추세를 보이기 때문에 월별 탑승객 수를 예측하기 위하여 가변수 회귀분석이 사용되었다. 가변수 회귀분석은 월별의 효과를 나타내기 위하여 다음과 같은 모델을 사용한다.

$$y_t = TR_t + SN_t + \varepsilon_t \quad (5)$$

$$y_t = \beta_0 + \beta_1 t + \beta_1 X_{s1,t} + \beta_2 X_{s2,t} \cdots \beta_p X_{s11,t} + \varepsilon_t \quad (6)$$

여기서 y_t 는 시간주기 t 에서 탑승객 수이고 TR_t 는 시간주기 t 에서 추세, SN_t 는 시간주기 t 에서 계절요인, ε_t 는 시간주기 t 에서 오차항, $\beta_0, \beta_1, \beta_2 \cdots \beta_p$ 는 알려지지 않은 모수, $X_{s1,t}, X_{s2,t}, \cdots X_{s11,t}$ 는 가변수이다. 예를 들면 시간주기 t 가 1월이면 $X_{s1,t}=1$ 이고 1월이 아닌 달은 $X_{s1,t}=0$ 이다.

3.3 Winters' 지수평활법

지수평활법은 시계열을 평활하게 예측하는데 사용되는 방법으로 근본적으로 먼 과거의 관측값들이 최근의 관측값보다는 중요하지 않다는 가정을 바탕으로 하고 있으며 가중값을 현재의 예측시점에서 멀어질수록 지수적으로 감소한다. 본 연구에서 사용된 데이터 세트가 강한 계절변동을 나타내기에 계절성 데이터에 적합한 Winters' 지수평활법이 사용되었다. Winters'방법의 자세한 수학적 분석은 Bowerman과 OComell[6]을 참조할 수 있다.

3.4 뉴럴네트워크 구조(architecture)

본 연구에서는 백프로퍼게이션이 명확한 수학적 논리를 기초로 하고 있고 여러 가지 응용 성공 사례를 가지고 있기에 뉴럴네트워크를 훈련시키기 위한 알고리즘으로 채택하였다. 입력과 출력변수의 똑같은 수가 뉴럴네트워크의 입력과 출력의 뉴런의 수로 결정이 된다. 즉 12개의 입력 뉴런(11가변수와 1개 시간변수)과 1개의 출력뉴런(탑승객 수)을 가진 뉴럴네트워크모델이 사용되었다. 11개 가변수는 월별 계절효과와 1개의 시간변수는 추세효과를 나타낸다. 단일 은닉층에서의 은닉뉴런의 수를 결정하기 위하여 12, 15, 17 은닉뉴런을 가진 뉴럴네트워크로 훈련데이터 세트를 제공함으로써 최적의 은닉뉴런의 수를 결정하였다. 최적의 은닉뉴런의 수를 결정하기 위하여 가장 적은 평균절대편차를 가진 은닉뉴런의 수를 채택하였다. 시그모이드 활성화함수가 은닉과 출력뉴런에 적용되었다. 안정된 학습을 위하여 학습속도를 0.05, 모멘텀을 0.5로 결정되었다. 훈련 후에 검증데이터세트가 뉴럴네트워크에 제시되었고 예측 결과가 제공되었다.

4. 실증결과분석

표2는 1991년 1월부터 1993년 12월 동안 검증데이터를 가지고 3개의 예측모델의 평균절대편차를 나타낸다. 훈련데이터세트가 3년일 때 정확도라는 기준에서는 뉴럴네트워크의 평균 MAD는 11302로 가장 좋은 예측모델임을 보여주었다. 또한 F-검증은 검증데이터 세트의 평균 MAD를 기준으

로 뉴럴네트워크와 회귀분석모델은 95%의 유의수준에서 중요하게 다름을 보여주었다. 훈련데이터 세트가 6년일 경우, 정확도를 기준으로 뉴럴네트워크의 평균 MAD는 10683으로 두 번째로 좋은 예측모델로 나타내었으나 F-검증에서는 95%의 유의수준에서 뉴럴네트워크의 수행이 다른 2개의 통계모델의 수행과는 중요하게 다르지 않았음을 보여주었다. 훈련데이터세트가 9년일 경우 뉴럴네트워크가 10628로 가장 낮은 평균 MAD를 보여주어 가장 정확한 예측모델이었고 F-검증에서는 뉴럴네트워크가 95%의 유의수준에서 회귀분석과 지수평활법보다 우수한 수행을 보여주었음이 입증되었다. 실증결과로부터 유의할 사항은 훈련표본크기가 3년에서 6년, 9년으로 증가될 때 뉴럴네트워크의 평균 MAD값은 11302에서 10683, 10628로 감소되었다.

<표 2> MAD를 기준으로 예측모델들의 수행비교

훈련데이터 크기	예측모델	1991	1992	1993	평균 MAD
3년	회귀분석	23799	11092	12354	15748*
	지수평활법	9341	14122	16095	13186
	뉴럴네트워크	13645	10054	10206	11302
6년	회귀분석	11712	14027	10845	12195
	지수평활법	6502	12804	12960	10755
	뉴럴네트워크	12715	9773	9562	10683
9년	회귀분석	14363	16296	14727	15129*
	지수평활법	12044	12441	18469	14318*
	뉴럴네트워크	11807	9542	10534	10628

註: 평균 MAD=(1991MAD+1992MAD+1993MAD)/3

* 95%의 유의수준에서 뉴럴네트워크가 다름

5. 결론

한국과 미국간 탑승객 수 예측을 위하여 뉴럴네트워크가 소개되었고 기존 시계열 통계분석과 비교되었다. 실증분석을 통하여 뉴럴네트워크의 예측능력이 전통적인 통계모델보다 우수함이 입증되었다. 특히 정확도라는 측면에서 뉴럴네트워크는 아주 뛰어난 수행을 보여주었다. 정확도는 항공기 운영계획을 위하여 중요한 관심사이므로 뉴럴네트워크는 의사결정자들에게 하나의 유용한 분석도구가 될 수 있다.

뉴럴네트워크는 기존의 통계분석보다 뛰어난 예측능력을 보여주었다. 이 새로운 테크닉은 배우기가 쉽고 상업용 뉴럴네트워크 소프트웨어에 의하여 쉽게 사용할 수 있다. 그러므로 항공기 의사결정자들은 탑승객 수의 예측을 위하여 뉴럴네트워크의 사용을 적극적으로 권장해야 한다.

참고문헌

1. B.L. Bowerman and R.T. O'Connell. *Forecasting and Time Series*, Duxbury Press, Belmont, CA, 379-431 (1993).
2. L.I. Burke. "Introduction to Artificial Neural Systems for Pattern Recognition," *Computers & Operations Research* 18, 211-220 (1991).
3. I. Chang, S. Rapiraju, M. Whiteside, and G. Hwang. "A Neural Network to Time Series Forecasting," *Proceedings of the Decision Sciences Institute* 3, 1716-1718 (1991)
4. N.R. Draper and H. Smith. *Applied Regression Analysis*, John Wiley & Sons, Reading, NY, 1-176 (1981).
5. K.A. Duliba. "Contrasting Neural Nets with Regression in Predicting Performance in the Transportation Industry," *Proceedings of the Annual IEEE International Conference on System Sciences* 4, 163-170 (1991).
6. S. Dutta and S. Shekhar. "Bond Rating: A Non-Conservative Application of Neural Networks," *Proceedings of the Annual IEEE International Conference on Neural Networks*, 443-450 (1988).
7. D. Hawley, J. Johnson, and D. Raina. "Artificial Neural Systems: A New Tool for Financial Decision-Making," *Financial Analysts Journal* 46, 63-72 (1990).
8. W.L. Huth and S.E. Eriksen. "Airline Traffic Forecasting Using Deterministic and Stochastic Time Series Decomposition," *The Logistics and Transportation Review* 23, 401-409 (1987).
9. K. Knight. "Connectionist Ideas and Algorithms," *Communications of the ACM* 33, 59-74 (1990).
10. W. Mendenhall and A. Sincich. *A Second Course in Business Statistics: Regression Analysis*, Dellen Publishing Co, Reading, CA, 2-246 (1993).
11. L. Marquez, T. Hill, R. Worthley, and W. Remus. "Neural Network Models as an Alternative to Regression," *Proceedings of the Annual IEEE International Conference on System Sciences* 4, 129-135 (1991).
12. M. Nelson and W.T. Illingworth. *A Practical Guide to Neural Nets*, Addison-Wesley, Reading, MA, 36-150 (1991).
13. D.B. Parker. *Learning Logic*, Technical Report TR-47, Center for Computational

Research in Economics and Management Science, MIT, Cambridge, MA, 1985.

14. D. Rumelhart, G.E. Hinton, and R.J. Williams. "Learning Representations by Back-Propagation Errors," *Nature* 323, 533-536 (1986).
15. D. Rumelhart, G.E. Hinton, and R.J. Williams. "Learning Internal Representations by Error Propagation," In D. Rumelhart, J. L. McClelland (Eds), *Parallel Distributed Processing*, MIT Press, Reading, MA, 318-362 (1986).
16. Volpe National Transportation Systems Center, Center for Transportation Information, DTS-44, Cambridge, MA, 02142.
17. P.D. Wasserman. *Neural Computing*, Van Nostrand Reinhold, Reading, NY, 1989.
18. P. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, Ph.D. Thesis, Harvard University, 1974.