M2 경쟁자료와 신경망[1]

M2 Competition Data and Neural Networks

지 원 철
홍익대학교 산업공학과
jhee@wow.hongik.ac.kr

<Abstract>

This paper is to evaluate the prediction accuracy of neural networks using the famous time series database. The second Makridakis Competition Data, which is abbreviated as M2 and consists of 29 time series, is used to test the forecasting ability of Multi-Layered Perceptron (MLP) and Cascade Correlation Network (CCN). The prediction accuracy of neural networks is compared with that of ARIMA model in terms of mean absolute percentage error (MAPE). MLP outperforms ARIMA in terms of fitting and forecasting accuracy, but CCN does not ourperforms the forecasting accuracy in spite of its excellence in fitting accuracy. The results shows that CCN can be a promising tool for forecasting if the learning algorithm of CCN is improved.

1. 서 론

근대적 시계열분석 모형의 개발은 1927년 Yule이 태양 흑점의 연간변동을 예측하기 위하여 사용한 자기회귀(Autoregressive, AR) 모형으로부터 비롯되었으며, 이후 반세기 동안 선형시스템에 의한 시계열 분석이 주류를 이루어 왔다. 하지만, ARIMA모형[Box & Jenkins, 1976]에 의해 대표되어지는 선형시스템에 의해 설명하기 어려운 현상들이 많다는 것을 인식하기 시작하면서 새로운 접근방법들이 개발되기 시작하였으며, 관련 연구들은 크게 두 가지로 분류할 수 있다. 첫 번째는 시계열의 생성과정이 확정적 모형에 의해 설명되어질 수 있다는 가정하에 동적 시스템 이론에 근거하여 시간지연(time-delay)을 포함하는 상태공간(state space) 재구성에 의해 시계열 생성과정을 모형화하려는 것이다[Harvey, 1984]. 두 번째는 신경망으로 대표되는 기계학습(Machine Learning) 분야의 발전에 영향을 받은 것으로 인공지능 연구의 관심이 규칙기반시스템으로부터 지식획득 문제의 해결을 위한 자료중심의 분석방법으로 옮겨감에 따라 시계열 분석에의 응용이 가능해졌다. 특히 1980년대 말에 들어와서는 비선형 모형에 대한 관심이 높아졌으며, 혼돈이론, 신경망 및 통계학 분야에서 개발된 다양한 기법들이 현실 세계의 비선형성을 효과적으로 다룰 수 있는 수단을 제공할 것으로 믿었다[De Gooijer & Kumar, 1992].

1) 본 연구는 과학재단의 연구비(과제번호: 951-0905-115-2) 지원을 받았음

하지만, 시계열 분석을 위한 새로운 접근방법들이 상당한 기대를 받았음에도 불구하고 예측성과에 있어서는 상당한 논란이 있어온 것을 부정할 수 없다. 이는 시계열 분석이 다양한 학문분야에서 연구되고 있으며 새로운 예측알고리즘의 개발이 비교적 쉬운 반면 해당 알고리즘의 예측성과 및 기존 알고리즘과의 연관성에 대한 정확한 평가는 쉽지 않았기 때문이다. 다시 말해 관련 연구성과들을 정확히 평가할 수 있는 수단의 부족으로 인하여 관련 문헌의 방대함에도 불구하고 연구결과의 신뢰성 문제 등으로 인하여 과거의 연구결과들을 바탕으로 새로운 연구를 시작하기 어려운 상황이었다.

이와 같은 문제점을 해결하기 위하여 Makridakis *et al*[1982, 1993]는 시계열 데이터베이스를 구축한 후 예측의 정확성을 경쟁하는 대회를 개최하였다. 이 대회는 흔히 Makrisdakis Competition으로 알려져 있으며 현재까지 시행된 세 차례의 대회를 구분하기 위해 M1, M2, M3라고 각각 부른다. 또, 신경망을 이용한 시계열 분석 연구를 해오던 Weigend도 Santa Fe 연구소와 NASA의 지원하에 같은 성격의 대회를 개최하였는데 사용된 시계열 데이터베이스를 Santa Fe 자료라고 부른다[Weigend & Gershenfeld, 1994]. 두 대회의 공통적 특징은 시계열 예측을 위한 공통의 시계열 데이터베이스를 구축한 후 예측방법에는 제한을 두지 않음으로써, 자연스럽게 여러 예측방법론들이 비교되도록 하여 시계열 연구자들에게 자신의 연구결과를 객관적으로 검증할 수 있는 기회를 제공하였다는 점이다. 두 시계열 데이터베이스는 단일변수 시계열분석(Univariate Time Series Analysis) 연구에 있어 중요한 벤치마크 자료로 활용되고 있다.

본 연구의 목적은 시계열 분석에 있어 신경망의 유용성을 검증하고자 하는 것이다. 이미 신경망을 이용한 시계열 분석은 새로운 연구가 아니지만 앞서 지적한 바와 같이 신경망의 예측성과에 대한 논란이 존재하므로 M2 경쟁자료를 사용하여 객관적인 평가를 하고자 한다. 과거의 연구에서 주로 사용되었던 신경망 모형들은 다계층 퍼셉트론(Multi-Layered Perceptron, MLP)이었는데, 본 연구에서는 최근 관심을 모으고 있는 Cascade Correlation Network(CCN) [Falhman & Lebiere, 1990]을 사용한 시계열 분석을 시도한다. 신경망을 이용한 시계열 분석방법들의 예측성과를 전통적 선형모형인 ARIMA 모형과 비교평가하기 위하여 AUTOBOX V3.0[AFA Inc., 1991]을 사용하였다.

2. 시계열 예측을 위한 신경망 모형

2.1. 신경망과 예측작업

신경망을 시계열의 생성과정에 대한 추정함수로 처음 사용한 연구는 Lapedes & Farber[1987]가 백프로파게이션 알고리즘에 의해 학습된 MLP가 기관측치들에의 적합은 물론 예측에 있어서도 기존의 계량적 기법들 보다 우수함을 보인 이후 시작되어 여러 연구들에 의해 뒷받침되었다[Werbos, 1988; Weigend *et al*, 1990,1991; Sharda & Patil, 1992; Jhee & Lee, 1993; Weigend &

Gershenfeld, 1994; Jhee & Shaw, 1996]. 하지만, White [1988], Fishwick[1989] 및 Ripley[1993] 등은 기존의 통계적 방법들 보다 신경망이 더 좋은 예측성과를 보인다고 할 수 없다는 연구결과를 발표하였다. 이와 같이 엇갈리는 연구결과들에 대해 Chatfield[1993]는 신경망이 새로운 유력한 예측 수단임에 틀림없지만 체계적인 실증적 분석이 필요함을 강조하였다.

최근까지 MLP는 시계열 분석에 있어 표준적 신경망으로 사용되어 왔는데 이는 MLP가 비선형 회귀분석을 하는 것과 같은 효과가 있으며 만능 근사기(Universal Approximator)라는 주장이 나올 만큼 주어진 자료에의 적합 능력이 뛰어나며[Hornick *et al*, 1989], 실세계의 시계열들 중 상당량은 비선형 모형에 의해 효과적으로 분석될 수 있기 때문이다. 하지만, MLP의 예측성과에 대한 평가가 일치하지 않는 이유중의 하나는 너무 많은 파라메터를 갖는 비선형 모형의 경우 과잉적합(Overfitting)의 위험이 크다는 점이다. 신경망의 과잉적합 문제를 해결하기 위해 다음과 같은 방법들이 제시되어 왔다.

1) 학습된 MLP로부터 불필요한 가중치(Weight)들을 제거하는 방법.
2) 사전에 MLP내 일부 가중치들의 값을 고정시킴으로써(Invariance의 도입) MLP내의 추정되어야 할 파라메터(Free Parameter)의 수를 줄이는 방법.
3) MLP가 주어진 학습자료에 대해 충분히 일반화 될 수 있도록 효과적으로 학습자료를 확장하는 방법.
4) 시계열 자료를 철저히 분석하여 MLP 설계에 대한 아이디어를 얻는 방법.

첫 번째 방법인 가중치 제거(Weight Elimination 또는 Pruning) 방법들은 catastrophic effect를 피하기가 쉽지 않고, 두 번째 방법인 신경망의 구조에 Invariance를 도입하는 방법은 사전에 명확한 신경망 설계지침이 존재하여야 하나, 현재까지 이에 대한 연구결과 부족과 MLP 내부의 구조변화로 인한 결과를 예상하기 어렵다는 점에서 현실적용성이 떨어진다. 또 세 번째 방법인 학습자료의 충분한 획득의 경우 사회경제 시스템으로부터 얻어지는 많은 시계열들은 관측치가 적고, 실험에 의한 추가 관측치의 생성도 불가능하다는 점을 고려하면 완전한 해결책이라고는 할 수 없다.

최근 Geman *et al*[1992]은 MLP를 현실문제에 적용함에 있어 발생되는 문제를 Bias/Variance Dilemma라고 표현하였다. 즉 Variance의 제어에 초점을 맞출 경우 추정된 모형의 Bias가 커지고 Bias의 제어에 초점을 맞출 경우 Variance가 커지므로 양자간의 적절한 타협점을 찾는 것이 필요한데 MLP는 구조상 이와 같은 타협점을 찾기가 쉽지 않다는 것이다. Geman *et al*의 주장은 그동안 제시된 MLP관련 학습이론들에 문제가 있는 것이 아니라, 주어진 문제를 MLP를 사용하여 표현하는데 어려움이 있다는 것이다. 즉, 신경망을 이용한 문제해결에 있어서 적합한 신경망 모형의 선택 및 선택된 신경망의 적절한 구조 및 크기를 결정하는 것이 가장 중요한 문제라는 것이다. 시계열 예측을 위해 리커런트 신경망 모형을 이용하거나, ARMA 모형 선정과정에서 얻은 결과를 신경망의 설계에 응용하는 연구들이나[Jhee & Lee, 1993; Jhee & Shaw, 1996], 본 연구에서 새로운 구조의 신경망인 CCN을 사용한 것은 모두 이러한 노력의 일환으로 볼 수 있다.

주어진 시계열 $\{Z_t,\ t \in T\}$에 대한 예측치를 얻기 위해 반복적으로 단단계 예측치 (iterative one-step-ahead forecast)를 얻는 절차를 사용하였다. 즉 과거의 관측치 $\{Z_t,\ t \in T\}$를 이용하여 시계열의 생성과정을 모형화한 예측함수 f를 추정한 후, f를 이용하여 예측기간 L 동안의 예측치들 $\{\hat{Z}_{t+r},\ r \in L\}$을 차례로 얻어가는 것으로 수식으로 표현하면 다음과 같다.

$$\hat{Z}_{t+1} = f(Z_t, Z_{t-1}, \ldots)$$

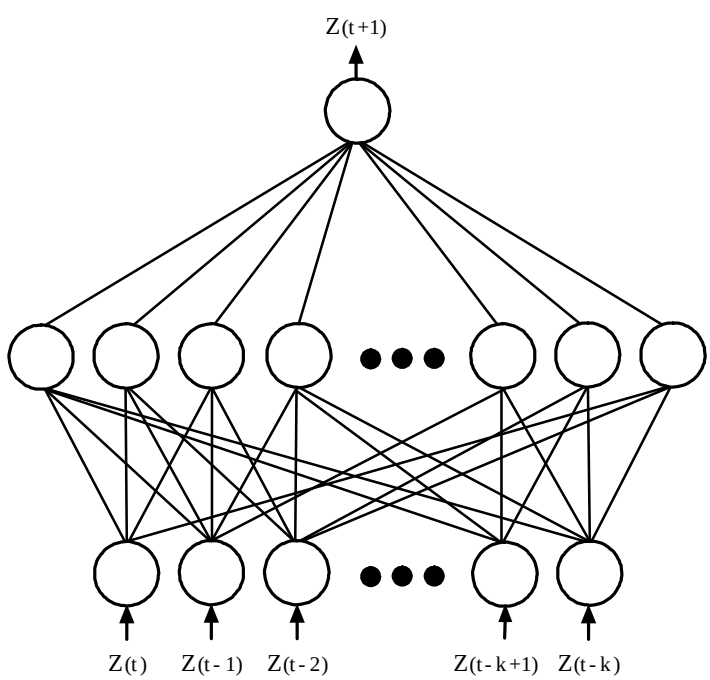
$$\hat{Z}_{t+2} = f(\hat{Z}_{t+1}, Z_t, Z_{t-1}, \ldots)$$

$$\vdots$$

$$\hat{Z}_{t+r} = f(\hat{Z}_{t+r-1}, \hat{Z}_{t+r-2}, \ldots, \hat{Z}_{t+1}, Z_t, Z_{t-1}, \ldots)$$

2.2. 다계층 퍼셉트론 (MLP)

시계열 예측에 사용되는 MLP의 기본 구조는 <그림 1>과 같다. <그림 1>의 MLP는 단단계 예측치(one-step-ahead forecast)를 얻기 위한 것으로 다단계(multi-step-ahead) 예측치를 원할 경우에는 출력노드의 간단한 변경으로 해결할 수 있다[Weigend & Gershenfeld, 1994; 지원철, 1995]. 본 연구에서는 하나의 은닉층만을 갖는 MLP를 사용했는데 이는 Hornick *et al*[1989]의 연구결과를 따른 것이다.



<그림 1> 단단계 시계열 예측치를 위한 MLP의 구조

MLP를 이용하여 좋은 예측성과를 얻기 위해서는 입력노드 및 은닉노드의 수를 적절히 결정하여

야 하는데, 입력노드 수의 결정은 MLP에만 국한되는 문제는 아니다. 하지만 은닉노드의 수를 결정하는 문제는 과거의 몇몇 연구에도 불구하고[Ash, 1989; Chauvin, 1989] 아직도 시행착오에 의존할 수밖에 없어 주어진 시계열에 적합한 MLP의 구조를 결정하기 위해서는 많은 시간과 노력이 필요하다. 이러한 어려움을 피하기 위하여, 본 연구에서는 새로운 신경망 모형 즉, Fahlman & Lebiere[1990]에 의해 제시된 CCN을 사용하여 은닉층의 노드 수를 결정하는데 따른 어려움을 해소하고자 한다.

2.3. Cascade Correlation Network (CCN)

Fahlman [1988]은 백프로퍼게이션 알고리즘의 학습속도가 느린 원인으로 Step-size 결정문제와 Moving Target 문제의 두 가지를 지적하고 있다. Step-size 결정문제는 표준 백프로퍼게이션 학습에 있어 전역오차함수에 대해 1차 부분도함수만을 사용하여 가중치의 갱신량을 계산하려하기 때문에 발생한다. 이 문제를 완화하기 위해 관성항을 사용하거나 2차 도함수를 사용하는 방법들이 많이 제시되었다[Wasserman, 1995]. Moving Target 문제는 MLP의 경우 신경망의 구조가 고정되어 있고 변화하는 입출력 패턴에 대해 MLP 내부의 가중치들이 모두 동시에 변화하게 되므로 생기는 문제이다. 즉 신경망 내부의 모든 노드들이 동시에 유용한 특성추출기(feature detector)가 되려하기 때문에 오히려 학습속도를 늦추고 경우에 따라서는 학습결과의 편차가 심해지는 현상을 보인다는 것이다. Fahlman과 Lebiere[1990]는 Step-size결정문제는 Quickprop 학습방법에 의해 해결하고 Moving Target 문제에 대해서는 새로운 신경망 구조를 제시함으로써 백프로퍼게이션 학습방법에 의한 MLP의 한계를 벗어나고자 했다.

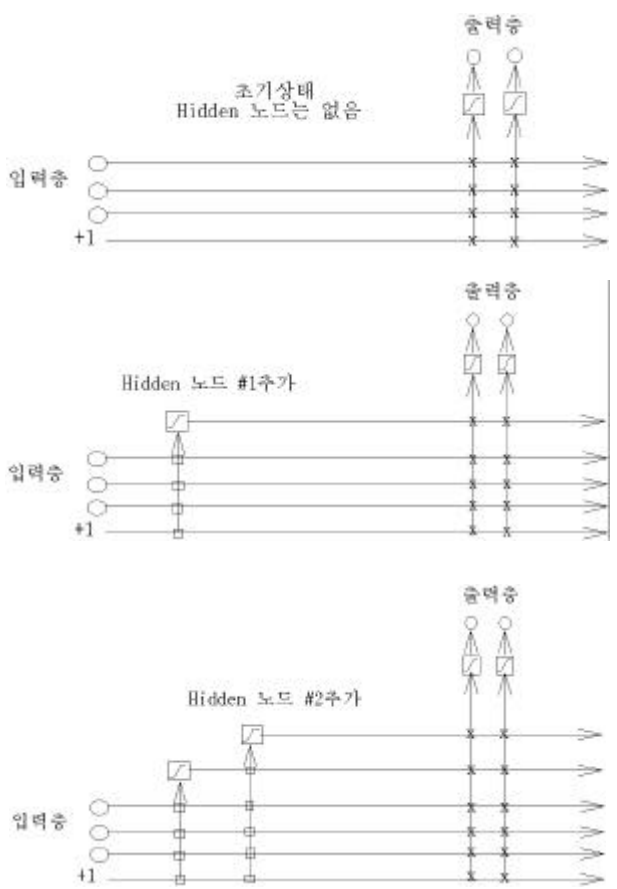
CCN의 구조는 <그림 2>와 같다. CCN의 학습은 은닉층의 노드들이 없는 상태에서 단순히 입출력 노드만을 가지고 시작한다. 학습을 진행시켜나가면서 오차의 감소가 더 이상 진전되지 않는 시점에서 일단 학습을 멈추고 해당 CCN의 성능을 평가하여 학습결과에 만족하면 학습을 종료하고, 만약 학습결과에 만족하지 못할 경우에는 은닉층의 노드들을 하나씩 순차적으로 증가시켜 가면서 학습과정을 반복한다. CCN을 시계열 예측에 이용한 연구로는 지 원철 등[1996]이 있다. 최근 CCN의 학습방법을 개선하고자 하는 연구가 많이 진행되고 있는데, 본 연구에서는 표준 학습방법을 사용하였다[Prechelt, 1997].

3. M2 경쟁자료

시계열분석을 위한 대표적인 벤치마크 데이터베이스로는 Makridakis 경쟁자료와 Santa Fe 자료를 들 수 있다. Santa Fe 자료는 총 6개의 시계열로 구성되어 있으며, 각 시계열내 관측치의 수도 1000개에서 3만개에 이를 정도를 충분한 관측치를 제공하고 있어 비선형 모형의 안정적 추정이

가능하다. 하지만 Santa Fe 자료가 사회경제 시스템으로부터 얻어진 시계열들이 아닌 반면에, Makridakis 경쟁자료는 모두 사회경제 시스템으로부터 얻은 시계열들로 구성되어 있다. 본 연구의 주된 목적이 관측치의 수가 적고 잡음이 심한 실세계의 시계열분석에 신경망을 적절히 사용할 수 있는 방안을 개발하는 것이므로 본 연구에서는 Makridakis 경쟁자료를 사용하였다.

Makridakis 경쟁대회의 목적은 다양한 단일변수 시계열 예측방법론들의 예측성능을 비교분석하기 위한 것으로 첫 번째 대회 결과는 1982년에 발표되어 학계에서 많은 논란을 불러일으켰으며 예측방법론의 발전에도 큰 공헌을 하였다[Makridakis *et al*, 1982]. 흔히 M1 경쟁자료로 불리는 시계열 데이터베이스는 총 1001개의 시계열들로 구성되어 있어 너무 방대하므로 101개로 간추린 데이터 베이스가 벤치마크 데이터로 많이 사용되어왔다. M1 경쟁자료 중 101개의 시계열에 대해 신경망을 이용한 분석도 몇 차례 보고된바 있다[Sharda & Patil, 1990; Jhee & Shaw, 1996].



<그림 2> Cascade Correlation Network의 구조 및 학습과정
(□ 표시 가중치는 고정되고 × 표시 가중치는 반복적으로 학습됨)

M2 경쟁대회는 M1 경쟁대회에 제출된 예측치들의 신뢰성에 대한 비판을 감안하여 시계열의 수를 대폭 축소하고 폐쇄적으로 개최되었으며 결과는 1993년에 발표되었다[Makridakis *et al*, 1993]. 최근 M3 대회가 열렸으나 아직 결과는 발표되지 않았다[Stewart, 1998]. 본 연구에서 사용될 M2 경쟁자료는 모두 29개의 시계열로 구성되어 있으며 23개의 시계열은 미국내의 서로 다른 산업에 속해있는 4개의 기업으로부터 직접 수집된 자료들이며 나머지 6개의 시계열은 거시경제 변수들에 관한 것이다. M2 경쟁자료를 관측기간과 관측치 수에 의해 정리한 것이 <표 1>이다.

<표 1> M2 경쟁자료의 분류

시계열 번호	관측치 수	관측기간
m001, m002, m003, m004, m005, m006, m020, m021, m022, m023, m024, m025	82	월별
m013, m014, m015, m016, m017, m018, m026, m027	94	
m019	45	
m028, m029,	237	
m007, m008, m009, m010, m011, m012,	167	분기별

4. 예측성과의 비교분석

4.1. 신경망의 학습조건의 설정

본 연구에서 예측함수 f를 추정하기 위한 사용된 신경망 모형들은 MLP와 CCN이며, 신경망의 예측성과를 비교하기위한 기본 모형으로 ARIMA이 사용되었다. 신경망의 예측성과를 결정하는 요인으로는 신경망의 구조, 학습방법 및 학습 파라메터들과 시계열의 생성과정 모형화에 사용될 입력공간의 크기, 즉 신경망 입력노드의 수 등 다양하다. 특히 신경망의 과잉적합 위험을 회피하기 위해서는 적절한 입력공간의 크기를 결정하는 것이 중요하기 때문에, 본 연구에서는 AUTOBOX 3.0을 이용한 ARIMA 모형화의 결과를 사용하였는데 주어진 시계열이 ARIMA(p,d,q)를 따를 경우 $Max(p+d, q+1)$를 신경망 입력노드의 수로 사용하였다[지원철, 1995; Jhee & Shaw, 1996].

신경망의 은닉노드 수 결정문제는 CCN의 경우 자동으로 결정되지만, MLP의 경우에는 사용자가 최종 결정하여야 한다. <그림 1>과 같이 MLP를 바로 예측함수로 사용할 경우 은닉층에 입력노드의 수와 같은 수의 노드를 사용하였다.

본 연구에서 사용된 모든 MLP들은 하나의 은닉층만을 가지고 있으며 은닉층과 출력층의 전이함

수로는 시그모이드 함수가 사용되었으며, 신경망의 학습을 위해서 시계열들은 (0, 1) 사이의 값들을 갖도록 선형변환하였다. 시계열들의 일부를 신경망의 학습에서 제외하여 학습된 신경망의 예측력을 검증하기위해 사용하였는데, 마지막 1년간의 좌측치들을 검증자료로 사용하였다. 따라서, 월별자료는 12개, 분기자료는 4개의 자료가 신경망 학습에서 유보되었다. 학습멈춤규칙은 MLP의 경우 최대학습회수가 50,000 Epochs에 이르거나 MLP의 RMSE가 0.0001 보다 작을 경우가 사용되었으며, CCN의 경우에는 CCN의 학습시 학습정도의 파악을 위해 많이 사용되는 EI(Error Index)의 값이 0.05에 이르면 학습을 멈추었다. Noutput을 출력노드의 수, NPattern을 학습패턴의 수, Output을 출력노드의 출력값이라고 할 때 EI는 다음과 같이 계산된다.

$$ErrorIndex = \sqrt{\frac{TrueError}{NOutputValue \times StdDev}}$$

$$where, \quad TrueError = \sum_i (\text{예측값} - \text{목표출력 값})^2$$

$$NOutputValue = Noutput \times N\ Pattern$$

$$StdDev = \sqrt{NOutput \times \sum Output^2 - (\sum Output)^2}$$

신경망의 출력값들을 학습자료의 구성시에 사용되었던 선형변환의 역으로 재변환하여 시계열의 실제값들과 비교되었다. 예측성과를 측정하기 위하여 MAPE(Mean Absolute Percentage Error)가 사용되었는데, 예를 들어 예측기간 L기간 동안 얻어진 예측치의 정확도를 측정하기 위한 MAPE의 계산은 다음과 같다.

$$MAPE = \frac{1}{L} \sum_{r=1}^{L} \frac{| \hat{Z}_{t+r} - Z_{t+r} |}{Z_{t+r}} \times 100$$

4.2. 실험 결과

M2 경쟁자료에 대해 ARIMA, MLP 및 CCN의 예측성과를 비교한 것이 <표 2>로, 신경망의 학습결과인 시계열에의 적합도 및 검증자료에 대한 신경망의 예측성과를 MAPE에 의해 측정한 상세 결과를 나타내었다. <표 2>의 평균과 표준편차에서 괄호안의 수자는 m015 (시계열명 : INTERSAL)의 경우 MAPE값이 학습과 예측에서 모두 이상적으로 매우 큰 값을 가지므로 이를 제외하고 평균과 표준편차를 구한 것이다. <표 2>에서 신경망의 학습능력, 즉 시계열에의 적합도는 선형 모형인 ARIMA에 비해 우수함을 알 수 있다. MLP와 ARIMA 사이에는 큰 차이가 없었지만 CCN의 적합능력은 매우 뛰어나다. 하지만 예측능력에 있어서는 MLP가 가장 좋았지만 통계적으로 유의한 결과는 아니었다.

<표 2> M2 경쟁자료 예측 모형들의 분석 결과 (MAPE)

	ARIMA		MLP		CCN	
	Fitting	Forecasting	Fitting	Forecasting	Fitting	Forecasting
1	16.78	10.65	16.27	19.73	1.86	12.94
2	21.2	16.9	19.07	14.79	2.34	14.81
3	12.17	17.6	0.92	13.43	0.73	14.72
4	8.98	12.81	9.47	11.35	0.71	14.89
5	8.63	10.14	9.37	11.19	0.63	14.5
6	8.77	6.01	4.36	8.09	0.49	13.19
7	0.79	1.04	1.59	4.85	1.01	2.6
8	0.45	0.37	2.89	5.43	1.82	4.34
9	1.24	0.85	1.84	1.13	2.15	0.9
10	1.95	2.04	2.07	5.41	1.68	3.64
11	0.38	0.78	4.92	6.3	2.07	5.49
12	10.01	13.4	10.05	12.44	6.39	15.61
13	25.12	25.51	23.37	19.24	4.31	27.85
14	18.31	24.26	11.68	36.84	7.91	38.64
15	85.89	185.43	87.37	147.96	6.15	116
16	17.61	17.66	16.47	15.37	2.37	18.9
17	8.93	8.42	8.49	9.35	1.4	14.74
18	4.64	8.46	5.15	6.43	0.86	7.6
19	25.39	48.02	24.17	25.54	5.13	37.78
20	17.8	32.36	12.44	35.18	3.39	24.99
21	17.34	36.42	22.48	17.39	4.7	25.97
22	17.38	37.3	21.38	27.96	5.88	38.88
23	17.56	35.99	25.64	26.93	4.75	26.63
24	2.38	2.52	1.89	2.95	0.5	4.3
25	2.7	5.82	3.32	8.7	0.63	9.72
26	5.75	4.67	3.54	4.77	0.65	8.97
27	6.87	5.5	4.61	5.31	0.92	7.61
28	5.48	3.14	5.31	4.3	1.06	4.13
29	4.77	5.59	4.65	4.28	0.9	5.72
평균	12.94 (10.36)	19.98 (14.08)	12.58 (9.90)	17.68 (13.02)	2.53 (2.40)	18.48 (15.00)
표준편차	15.96 (7.75)	34.43 (13.39)	16.39 (8.00)	26.85 (9.81)	2.15 (2.08)	21.72 (11.17)

CCN이 시계열에의 뛰어난 적합능력을 보였음에도 예측성능에 있어서는 세가지 모형 중에서 가장 나빴다. 본 연구에서는 CCN의 학습을 EI에 의해 일률적으로 제어함으로써 시계열에 따라 과잉적합 문제가 발생한 것으로 추정된다. 이는 <표 3>에 나타낸 CCN의 학습결과 생성된 은닉노드의 수를 살펴보면 지나치게 많은 수의 은닉노드가 생성된 경우들을 볼 수 있는데 이들은 과잉적합되었을 가능성이 크다. 이와 같이 CCN의 학습과정은 MLP의 경우보다도 제어하기가 쉽지 않지만 예측결과들의 표준편차는 ARIMA에 비해 작으므로 앞으로 계속적 연구가 수행될 경우 안정적인 예측 모형으로 사용가능하다고 판단된다.

<표 3> CCN에 의해 생성된 은닉노드 수

시계열 번호	은닉 노드수	시계열 번호	은닉 노드수	시계열 번호	은닉 노드수	시계열 번호	은닉 노드수	시계열 번호	은닉 노드수
1	21	7	5	13	14	19	18	25	14
2	33	8	2	14	5	20	17	26	12
3	10	9	5	15	29	21	23	27	11
4	13	10	28	16	13	22	21	28	35
5	11	11	2	17	15	23	33	29	21
6	14	12	41	18	23	24	9		

5. 결 론

본 연구에서는 신경망의 예측성과를 객관적으로 평가하기 위하여 두 번째 실시된 Makridakis 경쟁자료를 사용하였다. 또, 다계층 퍼셉트론(MLP)가 가지는 한계를 고려하여 신경망의 구조상에 장점을 가지고 있는 Cascade Correlation Network을 M2 경쟁자료 분석에 사용하였으며, MLP, CCN 및 선형모형인 ARIMA의 예측성과를 비교하였다. 시계열에의 적합능력은 CCN이 가장 우수했지만 유보된 자료에 대한 예측성능의 검증에서는 MLP가 가장 좋았다. 이는 아직도 CCN를 학습시키는데 있어 경험이 축적되지 않은 상태이며 과잉적합을 피할 수 있는 방안이 모색되지 않았기 때문이다. 하지만 최근에 CCN의 학습 알고리즘에 대한 연구가 많이 이루어 지고 있으므로 이에 대한 체계적인 검토를 통해 시계열 분석에 응용한다면 CCN은 유력한 예측수단이 될 수 있다.

6. 참고 문헌

지 원철, (1995). 신경망을 이용한 시계열 분석: M1-Competition Data에 대한 예측성과 분석. *한국전문가시스템학회지*, 1(1), pp135-148.

지 원철, 박 시우, 신 현정, 신 홍섭, (1996). Cascade-Correlation Network를 이용한 종합주가지수 예측. *한국경영과학회/ 대한산업공학회 96 춘계 공동 학술대회 논문집*, pp745-748

AFA Inc., (1991). *AUTOBOX 3.0 Manual*, Hatboro, PA.

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis - Forecasting*. San Francisco:Holden-Day.

Chatfield, C. (1993). Neural Networks: Forecasting Breakthrough or Passing Fad? *Int'l Journal of Forecasting*, 9, 1-3.

De Gooijer, J.G. & Kumar, K. (1992). Some Recent Developments in Non-linear Time Series Modelling, Testing, and Forecasting. *Int'l Journal of Forecasting*, 8, 135-156.

Fahlman, S.E. & C. Lebiere, (1990). *The Cascade-Correlation Learning Architecture.* CMU-CS-90-100, Pittsburgh, School of Computer Science, Carnrgie Mellon University.

Fishwick, P. A. (1989). Neural Network Models in Simulation: A Comparison with Traditional Modeling Approaches. *Proceedings of the 1989 Winter Simulation Conferenc*e, 702-710.

Geman, S., E. Bienenstock, E. & R. Doursat, (1992). Neural Networks and Bias/Variance Dilemma. *Neural Computation*, 4, 1-58.

Harvey, A.C., (1984). A Unified View of Statistical Forecasting Procedures. *Journal of Forecasting*, 3, pp245-275.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer Feedforward Networks are Uinversal Approximators. *Neural Networks*, Vol. 2, 359-366.

Jhee, W. C. & J. K. Lee, (1993). Performance of Neural Networks in Managerial Forecasting. *Int'l Journal of Intelligent Systems in Accounting Finance & Management*, 2(1), 55-71.

Jhee, W. C. & M. Shaw, (1996). Time Series Prediction Using Minimally-Structured Neural Networks: An Empirical Test. Trippi & Turban, (ed) *Neural Networks in Finance and Investing*, 2nd, 1996. also appears in *Proceedings of World Congress on Neural Networks*, 1994 Int'l Neural network Society Annual Meeting, San Diego, CA., II266-271.

Lapedes, A., & Farber, R. (1987). Nonlinear Signal Processing Using Neural Networks: Prediction and System Modeling. *Los Almos National Laboratory Report, LA-UR-87-2662*.

Makridakis, S, et al. (1982). The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. *Journal of forecasting*, vol. 1, No. 2, pp259-312.

Makridakis, S, et al. (1993). The M2-Competition: A real time judgmentally based forecasting study. *Int'l Journal of forecasting*, vol. 9, 5-22.

Prechelt, L., (1997). Investigation of the CasCor Family of Learning Algorithm. *Neural Networks*, 10(5), pp885-896.

Ripley, B. D. (1993). Statistical Aspects of Neural Networks. *Proceedings of 1993 SemStat*, Chapman & Hall, London.

Sharda, R. & Patil, R. B. (1992). A Connectionist Approach to Time Series Prediction: An Empirical Test. *Journal of Intelligent Manufacturing*.

Stewart, G.A., (1998). Prophetable Software, *Byte*, April, pp97-101.

Wasserman, P.D. (1993). *Advanced Methods in Neural Computing*. VNR, New York

Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1990). Backpropagation, Weight-elimination and Time Series prediction. In Touretzky (ed) *Advances in Neural information Processing Systems*, pp105-116.

Weigend, A. S. & N. A. Gershenfeld, (1994). *Time Series Prediction: Forecasting the Future and Understanding the Past*. Preceedings Volume XV in the Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley Publishing Company.

White, H. (1988). Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns. *Int'l Joint Conference on Neural Networks*, II451-458.