

# 퍼지시스템과 신경 회로망의 예측성능 비교분석

## Comparison of Fuzzy System and Neural Network as Predictor

공 창욱, 김 인택

Changwook Kong, Intaek Kim

명지대학교 정보제어공학과

### ABSTRACT

본 논문에서는 비선형 시스템 해석 문제에 널리 이용되고 있는 퍼지 시스템(Fuzzy System)과 신경 회로망(MLPNN)의 성능을 평가하기 위해 비선형 예측기를 구성하였고 두 예측기를 비선형 시계열(Time Series) 예측 문제에 적용하여 두 예측기의 성능을 비교 분석하였다. 예측 실험을 위한 데이터로 Mackey-Glass 와 Lorenz 시계열을 사용하였다.

### I. 서론

1980년대에 들어오면서부터 관심을 모으기 시작한 퍼지 시스템과 신경 회로망은 수학적 모델링 과정 없이 비선형 시스템의 해석이 가능하다는 장점 때문에, 과거의 고전적 이론으로는 해결하기 어려웠던 문제들 다시 말해 수학적 모델링이 복잡하거나 불가능하던 비선형 시스템 문제들에 적용되어 우수한 결과를 보여주었다[4][5]. 특히, 이들 퍼지 시스템과 신경 회로망은 비선형 예측기 구현에 많이 사용되어지고 있다[1][2][3]. 이와 같이 비선형 시스템 문제의 해결에 퍼지 시스템과 신경 회로망이 공통적으로 적용되고 있지만 각기 다른 방식의 학습방법과 몇 가지 학습변수 값들의 선정 절차가 수반되어 두 비선형 해석기의 성능을 비교할 수 있는 기준을 마련하기가 난해한 것이 사실이다. 본 논문에서는 퍼지 시스템과 신경 회로망을 이용한 각각의 예측기를 구성하고 두 예측기의 성능을 비교분석하기 위해서 비선형 시계열 예측에 적용하여 보았다.

## II. 예측기의 구성

### II-I 퍼지 예측기의 구성

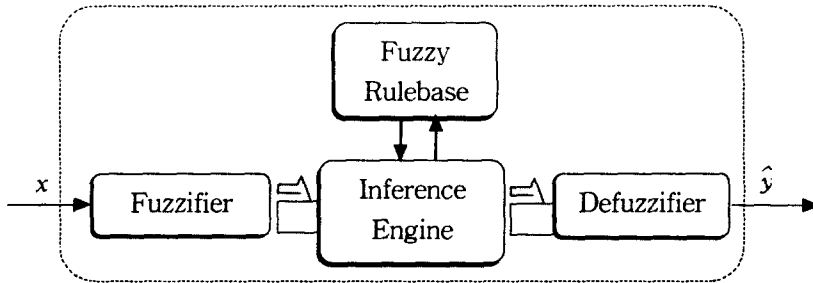


그림 1. 퍼지 예측기

퍼지 예측기는 과거의 제한된  $n$ 개의 데이터를 이용하여 일정시간 후의 값을 추정하도록 구성하였다(그림 1). 본 퍼지 시스템은 다음과 같은 함수로 표현될 수 있다.

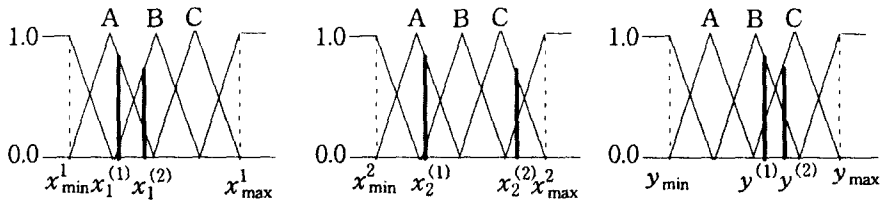
$$f: (x_1, x_2, \dots, x_n) \rightarrow y \quad (1)$$

$n$ 은 입력에 사용될 데이터의 개수이고,  $x_1, x_2, x_3, \dots, x_n$ 은 각각 퍼지 시스템의 입력과 출력 데이터이다. 퍼지 예측기에서  $y$ 는  $\hat{x}(n+1)$ 의 값을 가진다.

퍼지 예측기는 아래와 같은 순서에 따라 얻어진다.

i) 입력과 출력 분포 영역을 퍼지영역으로 나눈다.

$[x_{\min}^1, x_{\max}^1]$ ,  $[y_{\min}, y_{\max}]$ 과 아래와 같이 퍼지 영역을 정의하고  $(2N+1)$ 개의 퍼지 소속 함수를 정의한다.



ii) 학습 데이터로부터 퍼지 규칙을 생성시킨다.

다음과 같은 두 개의 데이터,  $(x_1^{(1)}, x_2^{(1)}; y^{(1)})$ ,  $(x_1^{(2)}, x_2^{(2)}; y^{(2)})$ 가 학습 데이터로 주어졌을 경우, 다음 순서에 따라 IF-THEN문의 퍼지 규칙 Rule1과 Rule2를 생성시킨다.

$$(x_1^{(1)}, x_2^{(1)}; y^{(1)}) \Rightarrow [(A, 0.8), (A, 0.7); (B, 0.8)]$$

⇒ Rule1 : IF  $x_1^{(1)}$  is A and  $x_2^{(1)}$  is A, THEN  $y^{(1)}$  is B

$(x_1^{(2)}, x_2^{(2)}; y^{(2)}) \Rightarrow [ (B,0.8), (C,0.6) ; (C,0.5) ]$

⇒ Rule2 : IF  $x_1^{(1)}$  is B and  $x_2^{(1)}$  is C, THEN  $y^{(1)}$  is C

iii) 생성된 각 퍼지 규칙의 선택 값을 구한다.

조밀한 분포의 데이터를 가지고 학습을 시킬 경우 전건부의 모든 퍼지 언어값(Fuzzy Linguistic Value)이 일치되는 여러 개의 퍼지규칙이 빈번히 발생하기 때문에 퍼지 규칙 선정에 혼돈을 가져올 수 있다. 이와 같은 퍼지 규칙 충돌의 문제를 효과적으로 대처하기 위한 여러 가지 방법들이 연구되어 지고 있고, 본 논문에서는 Max\_Min방법[1]을 사용하여 퍼지 규칙 충돌의 문제를 해결하였다. Max\_Min방법은 모순된 규칙의 퍼지 소속 함수값 중 최소 값을 선택하여(Max\_Min) 비교한 후 그 최소 값이 큰쪽(Max\_Min)을 선택하는 방법이다.

vi) 퍼지 규칙 베이스(Fuzzy Rule Base)를 구성한다

발생된 퍼지규칙들은 간단히 테이블의 형태로 표현할 수 있다. 전건부의 조건이 일치하는 다른 퍼지 규칙이 발생될 경우는 앞의 방법으로 나머지 퍼지규칙 제거하였다.

v) 비퍼지화(Defuzzification)

예측치를 구하기 위해 중앙 평균 비퍼지화(Center Average Defuzzification)공식을 사용하였다.

$$y = \frac{\sum_{i=1}^M \mu_{o_i} \tilde{y}^i}{\sum_{i=1}^M \mu_{o_i}} \quad (2)$$

## II-II 신경망 예측기의 구성

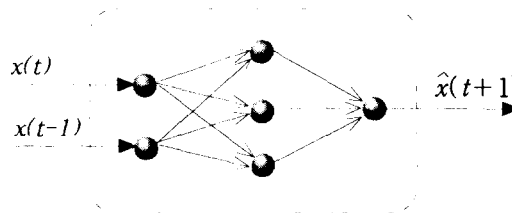


그림 2. 신경망 예측기

신경망 예측기로 널리 알려진 단방향 MLPNN( Multi-layer Perceptron Neural Network)을 사용하였고 학습 방법으로는 역전도(Error Backpropagation) 알고리즘을 사용하였다. 입력층(Input Layer)의 뉴런의 수는 입력조건에 따라 최대 다섯 개까지 변화될 수 있고 은닉 뉴런(Hidden Neuron)의 개수 역시 한 개부터 최대 다섯 개까지 증가시킬 수 있도록 구성하였다. 입력층, 출력층 그리고 은닉층(Hidden Layer)의 수는 각각 한 개로 제한하였다.

### III 시뮬레이션

#### III-I 실험 방법

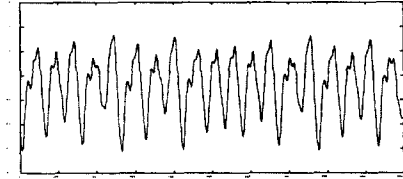


그림 3 Mackey-Glass 시계열

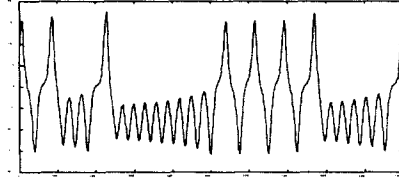


그림 4. Lorenz 시계열

본 실험에서는 Mackey-Glass 시계열과 Lorenz 시계열을 예측 대상 데이터로 사용하였다. Mackey-Glass 시계열과 Lorenz 시계열은 아래의 식(3)과 (4)와 같은 비선형 방정식으로 표현될 수 있다. 그림 3과 4는 각각의 방정식으로부터 발생된 1000개의 시계열 데이터들이다.

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \quad (3)$$

$$\frac{dx}{dt} = \sigma(y-x), \quad \frac{dy}{dt} = rx - v - xz, \quad \frac{dz}{dt} = sy - bz \quad (4)$$

퍼지와 신경망 예측기의 특성과 상관없이 기본적인 변수로 학습 구간의 길이 그리고 예측기 입력 데이터의 개수 등이 예측 성능에 커다란 영향을 미칠 수 있다. 본 실험에서는 다양한 결과를 얻기 위해 위의 두 가지 조건을 다음과 같이 가정하여 실험하였다. 총 1000개의 데이터를 갖는 Mackey-Glass 와 Lorenz 시계열 데이터 가운데 초기 100개, 초기 300개 그리고 초기 700개의 데이터를 각각 퍼지 예측기와 신경망 예측기의 학습 데이터 구간으로, 각각의 나머지 900개, 700개 그리고 300개의 데이터를 예측 구간으로 하였고, 입력 데이터의 조건은 연속되어지는 과거 2개의 데이터부터 최대 5개의 데이터까지를 예측기의 입력 조건으로 삼았다. 다음으로 각 예측기를 운영할 때 필요한 독립적인 변수를 조절하여 그 변화가 결과에 끼치는 영향을 살펴보았다. 먼저, 퍼지 예측기는 퍼지 소속 함수의 개수를 최대 21개까지 증가시켜 보았다. 신경망 예측기의 경우 은닉층의 뉴론(Hidden Neuron)의 개수를 최대 5개까지 증가시키고, 학습 수렴 속도를 결정하는 학습을 값들 [1E-03,1E-04,1E-05,1E-06,1E-07]로 변화시켜가며 그 관계를 살펴보았다. 신경망 예측기는 학습의 반복회수(Epoch)를 최대 10,000번으로 제한하였고, 퍼지 예측기의 경우 반복학습이 필요 없이 단 한 번의 탐색만으로 모든 학습을 마칠 수 있다. 예측 결과의 정확도와 성능 평가를 비교하기 위해서는 예측 수행 후 실제 치와 예측치 사이의 오차를 RMSE(Root Mean Square Error)로 측정하였다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x(i) - \hat{x}(i))^2} \quad (5)$$

### III-II 실험 결과 및 고찰

그림 6,7은 각각 MG와 Lorenz 데이터 700개의 학습 데이터를 가지고 나머지 300개의 데이터에 대한 예측을 수행한 후 퍼지 예측기와 신경망 예측기의 예측 오차를 RMSE값으로 도시한 것이다. 그림 6의 퍼지 예측기의 예측 오차를 보면 X축의 퍼지 소속함수의 개수와 Y축의 입력 데이터의 개수가 증가하는 방향으로 나란히 예측 오차가 줄어들고 있음을 알 수 있다. 그림 7을 살펴보면 신경망 예측기는 같은 실험에서 X축의 은닉층의 뉴론(Hidden Neuron)의 개수보다 Y축의 학습율의 변화에 더 민감함을 알 수 있다. 신경망 예측기에 대한 전체적 실험 결과에서 MG데이터는 학습 변수 값이 1E-04에 근접할수록, Lorenz 데이터는 1E-05에 가까워질수록 예측오차가 더욱 빠르게 감소하고 있음을 알 수 있었다. 그림 8은 1000개의 MG데이터 가운데 초기 700개의 데이터를 10,000회 학습시킨 후의 SSE(sum square error) 결과를 출력한 것이다. 학습율의 값 차이에 따라 학습오차가 다르게 감소되고 있는 모습을 볼 수 있다.

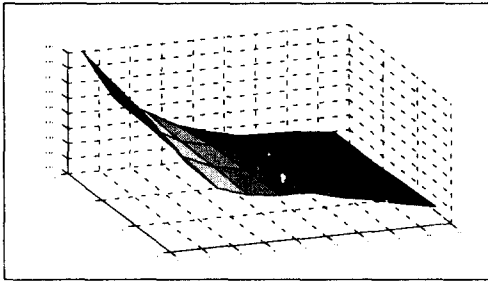


그림 6. 퍼지 예측기(Mackey-Glass)

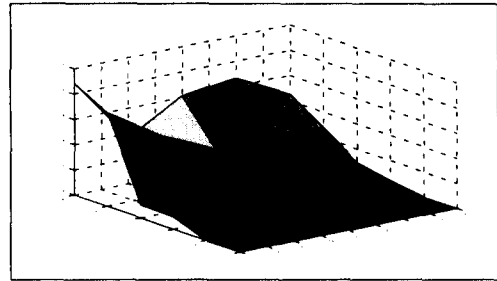


그림 7. 신경망 예측기(Lorenz)

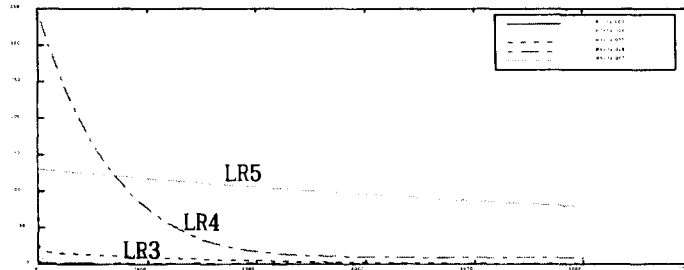


그림 8. 학습 변수의 변화에 따른 예측 오차의 감소

그림 9, 10은 신경망 예측기의 학습율이 예측 오차에 미치는 영향을 살펴보기 위해 Lorenz 시계열을 동일한 학습조건을 갖춘 두 예측기의 예측오차(RMSE)를 비교한 결과이다. 학습 변수가 1E-07인 그림 9에서는 신경망 예측기의 예측오차가 퍼지 예측기의 예측오차가 최소 값일 때보다 모든 조건에서 작았고, 그림 10에서는 학습율이 1E-08인 경우 퍼지 예측기의 성능이 더 우수하게

평가될 수 있음을 알 수 있다.

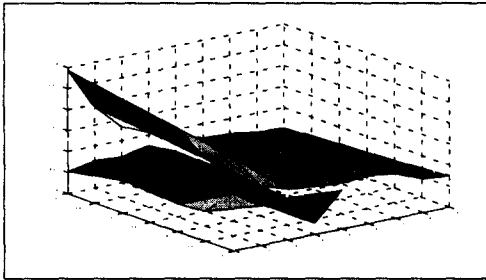


그림 9. 학습율=1E-07

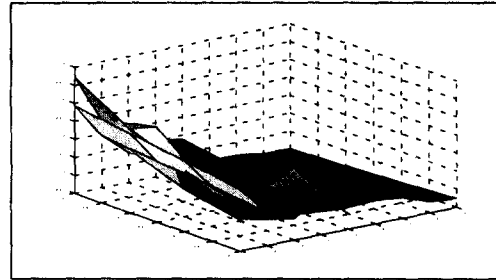


그림 10. 학습율=1E-08

### III. 결론

본 논문에서는 제안된 퍼지 시스템과 다층구조 신경 회로망으로 예측기를 구성하여 예측기의 성능과 예측 결과를 비교 분석하였다. 퍼지 예측기의 성능은 소속함수의 개수와 전건부 데이터의 개수 그리고 학습 데이터 구간의 길이에 따라 좌우된다. 신경망 예측기는 학습율 값의 설정이 성능에 가장 큰 변수로 작용하였다. 최적의 학습율 값의 설정은 경험적 실험결과에 의존할 수밖에 없다. 학습구간에 따른 비교에서 퍼지 예측기는 구간의 길에 따라 발생하는 퍼지 규칙의 수가 증감하기 때문에 퍼지 소속 함수의 개수와 더불어 퍼지 예측기의 예측성능을 결정하는 중요한 역할을 하였다. 예를 들어 학습 데이터 구간의 길이가 학습 데이터의 개수가 100개이고 퍼지 소속 함수의 개수가 비교적 많던 실험에서는 예측 실행 때 퍼지 규칙의 부재 문제가 발생하였다. 반면에 신경망 예측기의 경우 학습 데이터 구간의 길이에 따라 예측 오차가 감소하기는 하지만 큰 영향을 끼치지 않는았다.

### 참고 문헌

- [1] 김인택, 공창욱, "시계열 예측을 위한 퍼지 학습 알고리즘," 한국퍼지 및 지능 시스템 학회 논문지, Vol.7, No.3, pp.34-42, 1997.
- [2] E. Wan, "Time Series Prediction by using a Connectionist Network with Internal Delay Lines," *Time Series Prediction*, A. Weigend and N. Gershenfeld, Eds. Addison Wesley, 1994.
- [3] Wang, L. X. and J. M. Mendel, "Generating Fuzzy Rules by Learning from Examples," *IEEE Trans. on SMC*, vol.22, No.6, pp.1414-1427, Nov./Dec. 1992.
- [4] Wang, L. X., "Fuzzy systems are universal approximators," *Proc. IEEE International Conf. on Fuzzy Systems*, San Diego, pp.1163-1170, 1992.
- [5] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Network*, vol. 2, pp359-366, 1989.