

# 복합형 신경망을 이용한 시계열 예측 기법

한보경\*, 양홍민\*\*, 박소정\*, 김호준\*

\*한동대학교 전산전자공학부

\*\*한국과학기술원 전산학과

e-mail : hjkim@handong.edu

## A Time Series Prediction Technique using Hybrid Neural Networks

Bokyoung Han, Hongmin Yang, Sojeong Park, Hojoon Kim

School of Computer Science and Electrical Engineering, Handong Global University

Computer Science Department, KAIST

### 요 약

본 연구에서는 시계열예측 신경망기법에서 특징 선별기법을 적용하여 학습의 효율을 높이고 성능을 개선하는 방법론을 고찰한다. 순환구조 신경망을 사용하여 시계열 예측기를 구현하였으며, 효과적인 특징을 선별하기 위하여 FMM 신경망을 사용하여 특징의 연관도 요소를 산출하는 방법을 제시하였다. 모바일 결제시스템에서 실제 측정된 데이터를 사용하여 사용빈도 예측실험을 수행하였으며 그 결과를 통하여 제안된 기법의 유용성을 고찰하였다.

### 1. 서론

시계열 예측문제를 위한 신경망의 동작특성과 학습 기법에 관한 연구는 다양한 방법으로 시도되어 왔다. 최근의 연구에서 베이지안 필터와 퍼지신경망 모델을 이용한 시계열 예측기법[1]과 원자력발전소 운영 파라미터 예측문제를 대하여 동적 구조를 갖는 역전파 신경망 모델이 제안되었다[2]. 그런데 많은 응용문제에서 실제 환경에서 측정된 학습데이터는 환경변수 등의 요인으로 인해 비정상적인 특징 값이 포함될 수 있고, 이는 학습의 효율과 예측 성능을 현저하게 저하시킨다.

이에 본 연구에서는 시계열 예측 문제에서 학습데이터에 포함되는 불필요한 특징과 오류 데이터를 제거하기 위하여 변형된 구조의 FMM 신경망 모델[3]을 사용한 특징선별 기법을 제안한다. 신경망의 동작 특성에서 가중치 개념을 도입한 활성화 함수를 정의함으로써 학습데이터의 빈도 요소를 반영할 수 있을 뿐만 아니라 하이퍼박스 중첩영역의 모호성을 해결할 수 있도록 하였다. 또한 특징의 종류별로 연관도 요소(relevance factor)를 산출하는 방법을 제시하고, 이를 활용하는 특징선별 기법을 제안하였다.

제안된 시계열 예측 기법은 특징분석 모듈과 학습 모듈의 2 단계로 이루어지는 신경망 모델이다. 특징 분석 모듈에서는 FMM 신경망 모델을 사용하여 학습데이터에 포함되는 특징을 분석하여 효과적인 특징 집합을 생성한다. 학습 모듈은 순환구조 신경망을 사용하여 구현하였는데 입력 및 출력 노드에 대한 데이터 값 표현을 위하여 3 종류 퍼지집합의 멤버쉽함수를 정의하였다.

### 2. 시계열 예측을 위한 순환 신경망

본 연구에서는 시계열 예측의 응용문제를 위하여 다층 구조의 순환(recurrent) 신경망을 적용하였다. 일정 주기 단위로 측정된 과거 데이터로부터 일련의 특징값을 추출하여 입력데이터를 구성하며 3 종류의 퍼지집합으로 표현되는 출력값으로부터 시계열 데이터를 예측하게 된다.

측정된 특징값은 구간별 빈도 데이터의 최대값 및 최소값의 범위에 대하여 선형변환하여 정규화 한다. 정규화 과정을 거친 임의의 값에 대하여 3 종류의 퍼지집합에 대한 멤버쉽 값으로 각각 산출하여 표현하게 된다. 즉 특징값  $x$ 는 3 원소쌍  $(x^h, x^m, x^l)$ 으로 표현되며, 이들 각각은 식 (1), (2) 및 (3)으로 산출된다.

$$x^h = \begin{cases} 1.0 & \text{if}(R_{h_2} \leq x) \\ \frac{x - R_{h_1}}{R_{h_2} - R_{h_1}} & \text{if}(R_{h_1} \leq x < R_{h_2}) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$x^m = \begin{cases} 0 & \text{if}(x \leq R_{m_4}) \\ \frac{R_{m_4} - x}{R_{m_4} - R_{m_3}} & \text{if}(R_{m_3} \leq x < R_{m_4}) \\ 1.0 & \text{if}(R_{m_2} \leq x < R_{m_3}) \\ \frac{x - R_{m_1}}{R_{m_2} - R_{m_1}} & \text{if}(R_{m_1} \leq x < R_{m_2}) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$x' = \begin{cases} 0 & \text{if}(R_{i2} \leq x) \\ \frac{R_{i2} - x}{R_{i2} - R_{i1}} & \text{if}(R_{i1} \leq x < R_{i2}) \\ 1.0 & \text{otherwise} \end{cases} \quad (3)$$

### 3. 특징선별 기법

본 연구에서는 FMM 모델기반의 신경망으로부터 주어진 문제에 유용한 특징의 종류를 선별하는 방법을 제안하였다.  $j$  번째 하이퍼박스  $B$  는 입력패턴  $X_h = \{x_{h1}, x_{h2}, \dots, x_{hq}\}$  와 각 차원에서 특징값 범위의 상한값  $U_i = \{u_{i1}, u_{i2}, \dots, u_{iq}\}$  와 하한값  $V_i = \{v_{i1}, v_{i2}, \dots, v_{iq}\}$  및 이들을 적용한 활성화 함수  $b_i(X_h, U_i, V_i)$  로 구성된다. 하이퍼박스 활성화 함수의 정의는 식 (4)와 같다.

$$b_j(X_h) = \frac{1}{Z} \sum_{i=1}^n w_{ji} \cdot f(x_{hi}, I_{ji}) \quad (4)$$

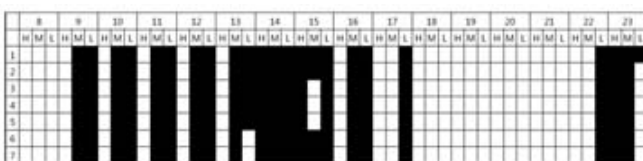
식에서  $X_h = \{x_{h1}, x_{h2}, \dots, x_{hn}\} \in I^n$  는  $h$  번째 입력패턴으로 총  $n$  개의 특징값들로 이루어진다. 또한  $I_{ji}$  는 하이퍼박스  $j$  에서  $i$  번째 차원에 대한 특징구간을 의미한다.  $Z$  는 정규화를 위한 상수이며,  $w_{ji}$  는  $i$  번째 특징과  $j$  째 하이퍼박스간의 가중치 요소이다. 학습된 FMM 신경망으로부터 특징과 패턴클래스간의 상호 연관도 요소(relevance factor)를 산출할 수 있다. 본 연구에서는  $i$  번째 특징과 패턴클래스  $k$  간의 연관도값  $RF(\cdot)$  를 식(5)와 같이 정의 한다.

$$RF(i, k) = \left( \frac{1}{N_k} \sum_{B_i \in C_k} w_{ji} - \frac{1}{N_B - N_k} \sum_{B_i \notin C_k} w_{ji} \right) / \sum_{B_i \in C_k} w_{ji} \quad (5)$$

식에서  $C_k$  는 클래스  $k$  에 속하는 하이퍼박스의 집합이며,  $N_B$  및  $N_k$  는 각각 전체 하이퍼박스의 개수 및 클래스  $k$  에 속하는 하이퍼박스의 개수를 의미한다. 이를 모든 패턴클래스에 적용하면, 주어진 응용에 대한 개별 특징의 유용성을 평가할 수 있다.

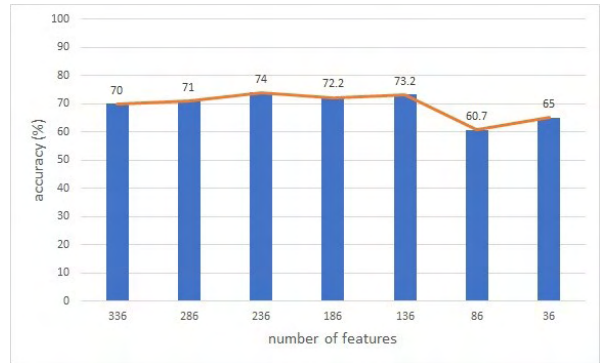
### 4. 실험 결과 및 고찰

제안한 이론에 대하여 모바일 결제 시스템 사용 빈도 데이터에 적용하였다. 순환신경망을 사용하여 시간대별로 데이터를 구분하였으며, 일일데이터를 기준으로 아침 8 시부터 24 시까지 1 시간 간격의 빈도를 예측하였다. 제안된 특징 선별 기법을 적용한 실험 결과를 (그림 1)에 보였다. 총 336 개의 특징에 대하여 연관도 요소를 산출하고 낮은 값의 특징 150 개를 제거한 결과이다. 그림에서 흑색으로 표시한 부분은 제거된 특징을 의미한다. 실험 결과 특징의 연관도 요소는 개별 특징별로 구분되는 것이 아니라 특정 시간대에 따라서 그 연관도가 달라지는 현상을 보였다.



(그림 1) 연관도 요소 평가 결과의 하위 150 개 특징을 제거한 결과

(그림 2)는 선별된 특징수에 따른 성능변화를 그래프로 보인 것이다. 효과적인 특징을 선택하여 적용하면 성능이 증가할 수 있으나, 반면 특징수를 지나치게 감소시키면 성능 저하가 발생함을 알 수 있다. 이 실험에서는 성능 저하 없이 총 336 개의 특징을 136 개로 줄일 수 있음을 보인다. 이는 성능저하 없이 60%까지의 특징수를 감소시킴으로써 학습 및 예측에 소요되는 계산량을 현저하게 개선할 수 있음을 보였다.



(그림 2) 선별된 특징수에 따른 예측 성능 변화

### 5. 결론

제안된 특징 선별 기법의 목표는 예측의 정확도를 저하시키지 않는 전제에서 특징수를 최소화하여 학습과 예측 과정의 효율을 개선함에 있다. 부수적으로 학습데이터에 포함되는 비정상적 데이터와 잘못된 특징값을 제거함으로써 부분적으로 예측 성능의 개선을 기대할 수 있게 한다. 실제 데이터를 사용한 실험 결과 특징수가 일정량 감소한 후에 오히려 정확도가 증가하는 현상이 관찰되었다. 이는 최초 특징 집합에 포함되었던 비정상적인 데이터에 영향으로 성능 저하가 발생함을 반영한다.

※ 이 논문은 미래창조과학부의 소프트웨어중심대학사업 (2017-0-00130)의 지원을 받아 수행된 연구임

### 참고문헌

- [1] Winkun Xiao, "Time series prediction using Bayesian filtering model and fuzzy neural networks," Optik, Vol.140, No. 1, pp.104-113, 2017.
- [2] Yong-kuo Liu, Fei Xie, Chun-li Xie, Min-jun Peng, Guohua Wu, Hong Xia, "Prediction of time series of NPP operating parameters using dynamic model based on BP neural network," Annals of Nuclear Energy, vol.85, no.1, pp.566-575, 2015.
- [3] Mohammed Falah Mohammed, Chee Peng Lim, "Improving the Fuzzy Min-Max neural network with a K-nearest hyperbox expansion rule for pattern classification," Applied Soft Computing, Vol.52, pp.135-145, 2017.