

# 온라인 시계열 자료를 위한 익스트림 러닝머신 적용의 최근 동향

## Recent Trends in the Application of Extreme Learning Machines for Online Time Series Data

윤여창<sup>†</sup>

우석대학교 정보보안학과

### 요약

익스트림 러닝머신은 다양한 방식의 예측 분야에서 주요 분석 방법을 제공하고 있다. 시계열 자료의 복잡한 패턴을 학습하고 잡음이 포함되어 있는 데이터거나 비선형인 경우에도 최적의 학습을 통하여 정확한 예측을 할 수 있다. 이 연구에서는 온라인 시계열 자료를 분석하는 도구로서 주로 연구되고 있는 기계학습 모형들의 최근 동향들을 기존 알고리즘을 이용한 응용 특성들과 함께 제시한다. 지속적이고 폭발적으로 발생하는 대규모 온라인 데이터를 효율적으로 학습시키기 위해서는 다양하게 진화 가능한 속성에서도 잘 수행될 수 있는 학습 기술이 필요하다. 따라서 이 연구를 통하여 시계열 예측 분야에서 빅데이터가 적용되는 최신 기계 학습 모형에 대한 포괄적인 개요를 살펴보고, 빅데이터에 대한 기계 학습의 주요 과제 중 하나인 온라인 데이터를 학습하는 최신 모형들의 일반적인 특성과 온라인 시계열 자료를 얼마나 효율적으로 학습하고 예측에 활용할 수 있는지에 대하여 논의하고 그 대안을 제시한다.

■ 중심어 : 익스트림 러닝머신, 온라인 데이터, 시계열 예측, 피드포워드 신경망, 순환 신경망

### Abstract

Extreme learning machines (ELMs) are a major analytical method in various prediction fields. ELMs can accurately predict even if the data contains noise or is nonlinear by learning the complex patterns of time series data through optimal learning. This study presents the recent trends of machine learning models that are mainly studied as tools for analyzing online time series data, along with the application characteristics using existing algorithms. In order to efficiently learn large-scale online data that is continuously and explosively generated, it is necessary to have a learning technology that can perform well even in properties that can evolve in various ways. Therefore, this study examines a comprehensive overview of the latest machine learning models applied to big data in the field of time series prediction, discusses the general characteristics of the latest models that learn online data, which is one of the major challenges of machine learning for big data, and how efficiently they can learn and use online time series data for prediction, and proposes alternatives.

■ Keyword : Extreme Learning Machine, Online Data, Time Series Forecasting, Feedforward Neural Network, Recurrent Neural Network

## I. 서론

시계열 자료는 시간 흐름에 대한 데이터 분석이 요구되는 분야에서 필요성이 증가하고 있다. 주식 시장, 강우량 및 토지 사용 패턴, 교통량, 디지털 신호 등과 같이 시간의 흐름에 따라 축적되는 데이터들은 시계열 분석의 중요한 대상 중의 하나다. 일반적으로 시계열 자료들은 반복적인 추세를 보이거나 또는 주기성이 없는 경향을 보일 수 있는데, 이들 자료는 예측, 분류, 함수 근사 및 탐색적 통계자료 분석에 주로 사용한다. 전형적인 시계열 자료 분석에 이용하는 응용 프로그램들은 경제 및 판매량 예측, 주식 시장 및 예산 추이 분석, 수익률 예측, 인구 조사 분석 등의 다양한 분야에 적용되고 있다. 또한 Han 등(2017)의 다변량 시계열 자료 예측을 위한 Gray 관계 분석 사용, O'Reilly 등(2017)의 일변량 및 다변량의 시계열 자료에 대한 매니폴드 학습 등과 같이 다양한 시계열 예측 알고리즘이 적용되고 있다.

그러나 이러한 알고리즘을 이용한 시계열 예측 접근은, 추세선을 연장하여 미래 시점에서의 상황을 예측하고자 시계열 자료에 맞는 외삽 적용 곡선을 광범위하게 사용하는 데 한계가 있었다. 전통적인 시계열 분석 방법의 하나인 자기회귀(AR) 모형 분석 기법은 시계열 자료 예측을 위한 다양한 통계적 접근 모형으로 확장되고 있다. 결과적으로 보다 정확한 시계열 예측 모형을 만들기 위해서는 시간 데이터를 가장 잘 표현할 수 있고, 시간 데이터에 영향을 미치는 다양한 변수들을 설명할 수 있어야 한다.

최근의 온라인 데이터들은 추세의 불규칙성과 비선형적 특성 그리고 대규모이면서 지속적이고 폭발적인 데이터 발생으로 인하여, 다루야 할 데이터양을 감당하지 못하는 경우가 많다. 발생된 데이터 수가 적다하더라도 실시간으로 발생한 데이터인 경우에는 적절한 분석 기법을 적용하는 데에 어려움이 많아 왔다. 또한 빅데이터에 대

한 기계 학습의 과제 중 하나는 지속적이고 폭발적으로 발생하는 대규모 온라인 데이터를 효율적으로 학습하는 것이다. 이러한 과제를 해결하려면 대규모로 발생하는 온라인 데이터와 진화 가능한 속성에서도 잘 수행되는 학습 알고리즘과 학습 기술이 필요하다.

MA, ARMA 및 ARIMA 등과 같은 모형들은 AR 모형의 발전과 더불어 다양한 시계열 자료의 모형 설정에 맞게 개선되고 있다. 그러나 온라인 시계열 모형의 효율성 높은 예측 결과를 위해서는 전통적인 통계 시계열 분석과 함께 이를 개선한 하이브리드 모형으로 진화가 필요하다. 이때 인공 신경망(Artificial Neural Network, 이하 ANN)의 익스트림 러닝머신(Extreme Learning Machine, 이하 ELM)과 더불어 유전자 알고리즘(Genetic Algorithm, 이하 GA)의 도입은 학습과 예측 효율성을 높이는데 도움이 된다.

## II. 데이터 분석

### 2.1 시계열 자료 분석

시계열 자료는 일정 기간 동안의 시간 흐름에 대하여 샘플링한 관측값이다. 이렇게 수집한 관측값들은 시간에 대한 추세에 영향을 미칠 수 있는 특징들을 가질 수 있다. 즉 일변량 및 다변량 시계열 자료는 시간의 흐름에 대한 데이터의 관측값이다. 시계열 자료는 일반적으로 다차원의 데이터이므로 차원 축소는 분석에 필수적이다. Wilson(2016), Gooijer와 Hyndman(2006)은 다변량 시계열 정보의 손실을 최소화하면서 기본 모형을 차원 축소하여 표현하는 몇 가지 방법을 논의하고 있다.

#### 2.1.1 PAA

단편화 근사 방식은 데이터를 시공간에 따라 분할한 다음 각 세그먼트에 대해 모형 적합 알고

리즘을 실행시킨다. 이 방법에는 각 세그먼트에 속한 시계열의 평균을 계수 값으로 근사시켜 결정하는 단편화 집계 근사 방법(Piecewise Aggregate Approximations, 이하 PAA), 적응형 단편화 상수 근사 방법(Adaptive Piecewise Constant Approximation) 그리고 다양한 특징들의 분할된 합(Segmented sum of various features) 등으로 발전하여 왔다(Yi와 Faloutsos, 2000). 이러한 근사화 기법들은 시계열 자료의 연속성을 유지하면서 동시에 차원 축소를 가능하게 한다.

### 2.1.2 SAX

기호 근사는 시계열 자료를 이산값(discrete)을 갖는 기호 변수로 변환시키는 시계열 표현 방법이다. 먼저 시계열 자료를 분할한 다음 분할된 단위에 대표적인 기호를 할당하여 실행한다. 주요 기호 표현 기술에는 시계열 자료를 포함하고 있는 대규모 데이터베이스에서 모호한 검색에 사용되는 기호 집계 근사(Symbol Aggregate approximation, 이하 SAX)가 있다. Lin 등(2007)의 연구와 Fayyad와 Irani(1993)의 연구에서 SAX는 단편화와 이산화 방법을 결합하는 대표적인 시계열 표현 방법으로, 연속 시계열의 차원 수를 줄이기 위해 PAA의 단편화 근사를 이용하여 이산 기호 형태로 변환한다. SAX는 PAA의 계산 능력을 더욱 향상시키고 다차원의 시간 데이터 차원을 효과적으로 줄인다. 반면에 각 세그먼트에 속한 시계열 값들은 평균화되기 때문에 시계열의 중요한 패턴들이 왜곡될 수 있는 단점이 있다.

## 2.2 시계열 예측 알고리즘

시계열 자료의 동적 특성을 더 잘 제어하기 위한 선형 모델로서 ARIMA 모형이 도입되면서 시계열 분석 알고리즘에 대한 접근 방향도 다양해졌다. 시계열 예측에 ARIMA 모형을 이용하면서 계산 편의성은 향상되었지만, George와 Mutalik (2019)에 의해 제기된 몇 가지 주요 문제를 해결

하는 데에는 실패했다. 첫째, 이 방법은 오프라인의 정적인 데이터에서 주로 동작하며, 온라인 시계열 자료의 대규모이면서 지속적이고 폭발적인 데이터 발생으로 인한 연속 데이터의 분석 평가 및 처리에 적용하기는 어렵다. 둘째, 일반적인 시계열 자료는 선형 모형을 따르는 경우가 많지 않고 비선형 모형으로 적합시키는 경우에도 어려움이 많다. 마지막으로, 시계열에 영향을 주는 외생 변수적 특성이 많이 있다. 이러한 특성들은 수학적 모델에 직접 측정하여 적합시키기 어려우며, 시간의 흐름에 대한 데이터 분석에 통합 적용하기도 어렵다.

따라서 이와 같은 비선형성 문제와 대규모로 발생하는 실시간 온라인 데이터 처리 문제를 해결하기 위하여 ANN을 적용하는 시계열 분석이 다양한 분야에서 이용되고 있다. ANN의 일반적인 특성은 비선형 데이터를 정확하고 일관된 모형으로 쉽게 매핑하여 설정할 수 있다. 그리고 ARIMA 기반의 모형과 결합하여 시계열 자료를 효과적으로 모델링하는 강력한 알고리즘을 제공하고 있다. 이들 모형에는 보다 크고 복잡한 구조에서도 운영할 수 있는 모형인 순환 신경망(Recurrent Neural Networks, 이하 RNN) 및 장단기 기억(Long Short-Term Memory, 이하 LSTM) 등이 있으며, 다양한 신경망 모형으로 결합 연구되고 있다. 또한 ELM 알고리즘은 Rosenblatt(1958)의 가장 단순한 퍼셉트론에서 시작된 신경망 학습의 일반적인 역전파 알고리즘(Backpropagation Algorithm, 이하 BPA)과는 다른 방법을 사용하여 단일 은닉층의 가중값만을 갱신하는 학습 알고리즘이다. ELM 모델은 Huang 등(2004)의 연구에서 시작하여, Tang 등(2016)에 이르러 데이터 학습 및 예측에 다층 퍼셉트론을 적용하면서 보다 효과적인 방법을 제공하기도 한다.

### III. 인공 신경망

많은 연구자들이 시계열 예측을 위해 ANN을 이용하는 것은, 전역 최적해로 근사시키는 일반적인 알고리즘을 사용할 수 있기 때문에 가능하다. 대부분의 신경망에서 가중값들은 BPA를 사용하여 업데이트한다. 확률적 기울기 감소 기반 알고리즘은 최적의 가중값을 찾기 위한 방법이다. 그러나 이 알고리즘의 주요 문제는 전역 최소값으로의 수렴 속도가 매우 느릴 수 있다는 것이다. Smith(1996)에서는 학습에 이용하는 은닉층의 수와 각 은닉층을 나타내는 뉴런의 수가 임의적이며, 이들의 효율적인 결합을 통해 시계열 모형의 비선형성을 해결하고 있다.

그러나 온라인 데이터에 적용할 수 있어야 한다는 원론적인 문제는 해결되지 않고 있다. 시계열 자료의 정확한 예측을 신속하게 제공하지 못하는 가장 큰 이유는, 이용할 데이터의 양이 폭발적으로 증가하는데 반하여 신경망 학습의 수렴 속도가 느리다는 데에 있다. 이 문제를 해결하기 위하여 현재까지 최첨단 신경망 알고리즘을 사용하는 몇 가지 주요 모형들이 지속적으로 연구되어 왔다[11,12,15-18].

#### 3.1 FNN

피드포워드 신경망(Feedforward Neural Network, 이하 FNN)이란 흐름이 단방향인 신경망을 말한다. 즉 입력 신호가 다음 층 즉 중간층으로 전달되고, 그 신호를 받은 층은 그 다음 층으로 전달된다. 즉 FNN에서는 BPA를 사용하지 않고 가중값의 업데이트가 이루어지므로 신경망은 한 방향으로만 작동한다. FNN은 구성이 단순하여 구조를 이해하기 쉽고, 따라서 많은 문제에 응용할 수 있다. FNN은 주어진 가중값으로 각 층을 계산 평가함으로써 최종의 예측을 제공하는 방식으로 동작한다. 그러나 시계열 자료를 잘 다루지

못한다는 단점뿐만 아니라 단순한 구조의 FNN에서는 시계열 데이터의 성질을 충분히 학습할 수 없다.

그러므로 FNN을 사용하는 최적의 방법으로는 다음 장에서 설명하는 ELM 모형의 적용 시에 가장 적합할 수 있다.

#### 3.2 RNN

RNN이란 어느 한 지점에서 시작한 학습이 시간을 지나 다시 원래 장소로 돌아오게 되고, 이 과정은 닫힌 경로 안에서 반복하는 순환 네트워크를 의미한다. 닫힌 경로를 통하여 데이터가 같은 경로를 반복 학습하면서, 데이터는 순환하고 정보는 끊임없이 갱신된다. 그리고 데이터가 순환하기 때문에 과거의 정보를 기억하는 동시에 최신 데이터로 갱신될 수 있다.

Kuo 등(1992), Li와 Lin(2016)의 연구에서는 RNN을 시계열 예측에서 적용하기 위하여 SCRFNN이라는 새로운 유형의 네트워크 아키텍처를 제안하여 카오스 시계열 예측을 한다. Qinghai와 Lin(2016)의 연구에서는 RNN 적용 분석을 다음과 같이 두 단계로 개선시키고 있다. 첫 번째 단계는 시계열 추세를 정확하게 모델링하는 동적 학습 단계이고, 두 번째 단계는 각 세그먼트의 오류 비용을 결정하는 Lyapunov 지수를 이용하여 복잡하고 예측 불가능한 카오스상태의 지수화 설정과 시간 지연 신경망(Time Delay Neural Network, 이하 TDNN)을 사용하는 단계이다. 신경망과 ARIMA 모형의 비교 연구인 Ho 등(2002)의 연구 결과에서는 RNN이 ARIMA 모형보다 더 좋은 성능을 나타내고 있음을 보여주고 있다. 다른 시간 내에서 유사한 입력 패턴의 행동 변화를 모형화할 수 있는 RNN의 기능은 네트워크를 더욱 강력하게 만들어준다. 네트워크의 가중값들은 일반적으로 BPA를 통해 업데이트된다. 감쇠된 피드백 가중값의 존재도 모형

성능에 영향을 준다. 이들 연구에 의하면 RNN은 단기간 시계열 자료를 이용한 예측인 경우에 가장 적합하며, 더욱 정확한 예측을 위해서는 감쇠 계수의 적절한 값의 선택에 따라 보다 큰 영향을 받는다.

### 3.3 LSTM

RNN은 자연어 처리나 시계열 자료와 같은 순차적 데이터를 처리하는 데 주로 사용되는 것으로 이전 시점의 정보를 은닉층에 저장하는 방식을 취한다. LSTM은 이전 정보를 기억하는 정도를 적절히 조절해서 입력값과 출력값 사이의 시점이 멀어질수록 이전 데이터가 점점 사라지는 기울기 소멸 문제(vanishing gradient problem)를 해결할 수 있다. LSTM은 기존 기울기 소멸 문제를 해결할 수 있다는 장점에 반하여, 과거 데이터의 제어 구조가 복잡하여 연산 속도가 느리다는 단점도 있다.

LSTM은 데이터의 장기 종속성을 학습할 수 있는 RNN이다. Gers 등(2002)의 연구에서는 단순 시계열 예측 알고리즘을 대체하기 위한 모형으로서 LSTM을 사용하고 있다. 여기에서 LSTM은 유사한 입력 패턴 간의 장기 유사성을 학습하면서 AR 모형으로 모델링한다. 따라서 LSTM은 TDNN 접근 방식과 비교할 때 순수한 AR 모형과 상대적으로 가깝다. 즉, Hou 등(2020)의 연구에서 LSTM 모델은 시계열 자료에 쉽게 액세스할 수 있으며, 더 나은 예측을 위해서 이전에 모델링한 함수에 시계열 자료를 적용한다. 그러나 Hou 등(2020)의 연구 결과에서 LSTM을 사용하면 시계열 자료의 변동에 대한 모델링은 확실히 향상되지만 과거의 중요한 이벤트에 영향을 받는 필수 속성에는 도움이 되지 않는다. 따라서 LSTM이 RNN의 TDNN 접근 방식보다는 더 잘 동작할 수 있지만 유의적 장점을 제공하지는 않는다고 할 수 있다.

## IV. 익스트림 러닝머신

FNN은 경사하강법 기반 학습, 표준 최적화 학습 그리고 최소 제곱 기준의 학습 등의 다양한 방법으로 학습을 하지만, 앞 장에서 살펴본 바와 같이 학습 속도가 일반적으로 매우 느린 기울기 기반 학습 알고리즘이라는 단점이 있다. 그러나 FNN은 신경망을 학습하는 데 광범위하게 사용되며, 네트워크의 모든 매개변수가 학습 알고리즘을 통하여 반복적으로 갱신되어 조정된다.

Tang 등(2016)의 연구는 단일 은닉 레이어 피드포워드 신경망(Single-hidden Layer Feedforward Neural Network, 이하 SLFNN)을 위한 개선된 ELM이라고 할 수 있는데, FNN 학습의 단점을 해결하기 위하여 입력 가중값을 무작위로 선택할 수 있게 한다. Huang 등(2011)의 연구에서 ELM은 분류, 회귀, 클러스터링, 희소 근사, 압축 및 특징 학습을 위한 FNN으로 분류된다. 이 신경망은 은닉 노드가 단일 또는 다중 레이어로 구성되고, 은닉 노드의 가중값들은 무작위로 할당되어 갱신되거나 상속될 수 있다. 대부분의 경우, 은닉 노드의 출력 가중값은 일반적으로 각 반복 단계에서 갱신 학습되며, 이는 본질적으로 선형 모형을 학습하는 것과 같다. Markowska-Kaczmar와 Kosturek(2021)의 연구에서는 상대적으로 작은 데이터 세트의 경우 ELM이 BPA로 훈련된 네트워크 학습 결과보다 우수하다는 것을 보여주었다. 그러나 까다로운 이미지 데이터 세트와 같이 대규모 입력 데이터인 경우에 ELM은 BPA로 훈련된 최신 네트워크에 비해 오히려 경쟁력이 없었다.

ELM이라는 이름은 단일 레이어 퍼셉트론으로 시작된 Huang 등(2004)의 연구로부터 시작한다. 이 연구의 ELM 알고리즘은 빠른 학습 속도와 최고의 일반화 성능을 제공한다. 또한 Tang 등(2016)에서 설명하는 ELM은 신경망을 학습시키는 또 다른 대안적 방법이며, 은닉층의 노드

수 변화에 한결같이 실행되는 일반화 SLFNN 네트워크로 동작한다. ELM은 FNN에서 최상의 일반화 네트워크를 제공하고 있다. 최근 들어 시계열 자료에 대한 ELM 기반 알고리즘의 적용이 여러 연구에서 적용되고 있다. Ahmad 등(2018)과 Wang 등(2019)과 같은 일부 연구에 의하면, ELM 모델은 좋은 일반화 성능을 낼 수 있으며 BPA를 사용하여 훈련된 네트워크보다 수천 배 빠르게 학습할 수 있다고 하며, 분류 및 회귀 응용에서 서포트 벡터 머신을 능가할 수 있다는 것도 밝히고 있다.

#### 4.1 OS-ELM

온라인 순차 익스트림 러닝머신(Online Sequential ELM, 이하 OS-ELM)은 BPA의 수렴 속도 문제를 해결하기 위한 방안으로 고려되었다. 이 모형은 Smith(1996)의 일반적인 초기 랜덤 발생값들을 이용하여 신경망의 초기 가중값을 발생시킨다. 가중값들의 갱신과 이들의 분석은 학습 과정 중에 새롭게 구성되는 방정식들을 사용하여 실행된다. 초기 가중값에 대한 OS-ELM의 높은 의존성은 이 모형의 최대 단점이라고 할 수 있다. 따라서 OS-ELM은 고차원의 초기 데이터 세트를 이용하여 안정적인 역행렬을 구할 수 있어야 한다. 상당한 양의 초기 데이터 값들을 사용하여 가중값 행렬을 초기화한다면, OS-ELM 모형은 시계열 자료를 예측하고 모델링하기 위해 정확하게 작동할 수 있다. OS-ELM 모형을 이용한 최근의 연구들은 초기 가중값의 선택 문제와 지역 최소값과의 포화 문제를 해결하는 방향으로 진행되고 있다.

또한 심층 신경망의 OS-ELM의 경우 적절한 학습률(learning rate)을 선택하는 것은 매우 어렵기 때문에, Bharath(2022)의 연구에서는 카오스 시계열 예측에 사용되는 FNN의 OS-ELM 성능 향상을 위하여 그리피스의 가변 학습률(Griffiths'

Variable Learning Rate)을 이용하고 있다. 여기서 학습률은 학습 데이터와 제공 오차 사이의 Griffiths의 상호 상관을 기반으로 다양하므로 시계열 자료를 더 잘 추적할 수 있다.

#### 4.2 OS-ELM의 후속 연구

빅데이터에 대한 기계 학습의 주요 과제 중 하나는 대규모이고 지속적이면서 폭발적으로 발생하는 온라인 데이터를 효율적으로 학습하는 것이다. 이러한 과제를 해결하려면 대규모 데이터와 함께 진화 가능한 속성에서도 잘 수행되는 학습 알고리즘과 학습 기술이 필요하다. OS-ELM은 더 많은 응용 분야에서 대규모 빅데이터를 계산하기 위한 실행 가능한 대체 기술로서 강력한 잠재력을 가지고 있다. 이러한 작업은 OS-ELM에서 가장 중요한 작업 중의 하나이다. Chang 등(2019)의 연구에서는 OS-ELM의 단점을 극복하기 위한 개선된 OS-ELM에 대한 최신 연구들로서 일부 전문성을 추가한 확장 버전에 대한 비교 연구들을 제시하고 있다.

OS-ELM의 후속 연구들에는 Sun 등(2007), George와 Mutalik(2019) 그리고 Rajendra와 Motupalli(2022) 등이 있다. Liem(2022)의 연구에서도 OS-ELM에서 가장 중요하고 뚜렷한 결함이라고 할 수 있는 초기 가중값 선택의 문제를 극복하기 위한 후속 연구가 진행되고 있다. OS-ELM의 다양한 변형으로서의 연구 개선 결과는 다음과 같다.

##### 4.2.1 ROS-ELM

정규화 OS-ELM (Regularized OS-ELM, 이하 ROS-ELM)은 대각 행렬 계수(diagonal loading factor)를 사용한다. ROS-ELM은 OS-ELM과 동일하게 재귀적 방식으로 동작하지만 향상된 수치적 속성으로 더 나은 성능을 제공한다. 그러나 Zhou 등(2014)의 연구에서는 OS-ELM과 비교할 때 학습 수렴의 속도에는 유의적 차이는 없었다.

#### 4.2.2 KOS-ELM

OS-ELM 모델은 재귀적인 최소제곱 알고리즘이라 할 수 있지만, KOS-ELM (Kernelized OS-ELM)은 OS-ELM의 커널화된 접근 모델이라고 할 수 있다. Zhou 등(2014) 그리고 Scardapane 등(2015)에 의해 연구되어 있는 커널화 접근은 OS-ELM의 목적을 기능 확장으로 수행하면서 학습의 적합 오류를 최소화하는 최적의 초기 가중값을 찾을 수 있게 하고, KOS-ELM 모형이 선형 또는 비선형 시계열 자료를 완벽하게 매핑할 수 있음을 보여주고 있다. 반면에 커널은 엄청난 양의 계산량이 요구된다. 따라서 KOS-ELM을 이용할 때 고려해야 하는 주요 문제는 비효율적인 계산량이다. KOS-ELM을 통하여 얻은 일반화 오류 및 학습 시간 측면을 비교할 때, Scardapane 등(2015)의 연구는 개선된 계산량을 보여주는 매우 효율적인 알고리즘을 제시하고 있으며, 경험적인 연구 결과에서는 일부 벤치마킹 데이터 세트에 대해서도 유의한 결과를 보여주고 있다.

#### 4.2.3 OSF-ELM

OSF-ELM (Online Sequential Fuzzy ELM)은 0차 및 1차 TSK (Takagi Sugeno Kang) 퍼지 추론 시스템(Fuzzy Inference System, 이하 FIS) 모형을 실행시키기 위하여 방사형 기저함수(Radial Basis Function, 이하 RBF)를 사용한다. Sun 등(2007)의 연구에서는 먼저 처리되는 데이터를 k-평균 클러스터링 방법으로 그룹화하여 일반적인 TSK 퍼지 추론 시스템을 개선한 버전이다. Matias 등(2015)은 TSK의 모형 접근 방식으로 연구한 결과로서 ELM을 일반화 모형으로 확장시킬 수 있다. Rong 등(2009)의 연구 결과는 FIS 모형을 SLFNN으로 실행시키면서 OS-ELM 모형에 대한 FIS 구현의 중요한 요소를 제시하고 있다. 퍼지 모델을 통하여 측정할 수 없는 외생적 요인에 대한 논의는 George와 Mutalik(2019)에서 논의된 결과와 같다.

#### 4.2.4 AWOS-ELM

시계열 문제를 예측하기 위한 새로운 적응형 가중값 온라인 순차 익스트림 학습기(Adaptive Weight Online Sequential ELM, 이하 AWOS-ELM)는 OS-ELM을 기반으로 시계열 문제를 학습한다. 온라인 애플리케이션에서 순차적으로 들어오는 데이터 집합은 일반적으로 다양한 신뢰 계수를 가지며, 신뢰 계수가 낮은 데이터 집합은 후속 교육 프로세스가 학습 과정을 잘못된 방향으로 이끄는 경향이 있다. Lu 등(2017)이 제안한 AWOS-ELM은 신뢰도 계수에 적응적으로 접근시킨다. 이에 따라 훈련 가중값을 결정함으로써 OS-ELM의 단점인 훈련 과정을 개선할 수 있다. 이 연구의 시계열 자료 예측에 대한 실험을 통해 AWOS-ELM 알고리즘이 OS-ELM 알고리즘보다 일반화 성능, 안정성 및 예측 능력에서 더 나은 성능을 보이는 것으로 확인되었다.

#### 4.2.5 PSTOS-ELM

온라인에서 새로운 학습 데이터를 업데이트하는 학습 모형이라고 할 수 있는, 점진적인 구조적 관용 온라인 순차 익스트림 학습 머신(Progressive Structural Tolerance Online Sequential ELM, 이하 PSTOS-ELM)은 Atsawaraungsuk 등(2023)에 의해서 연구되고 있다. 이들의 새로운 실험 모형인 PSTOS-ELM은 온라인 데이터 스트리밍이 있는 경우에 적합하며, 새로운 학습 데이터를 업데이트하는 점진적 학습을 통하여 강력한 정확도를 유지할 수 있음을 보여주고 있다. 이 연구에서는 새로운 클래스 데이터 사례에서 데이터를 재교육해야 하는 ELM 및 구조적 허용 오차 온라인 순차 익스트림 학습 머신(Structural Tolerance Online Sequential ELM, 이하 STOS-ELM)을 사용하여 데이터가 업데이트되는 동안에도 PSTOS-ELM의 예측 정확도가 상대적으로 높았다. 결과적으로 PSTOS-ELM은 ELM 및 STOS-ELM에 상응하는 정확성과 예측 일치성을 가지며 새로운

클래스 데이터를 즉시 점진적으로 갱신할 수 있음을 보여주고 있다.

#### 4.2.6 OSLA

온라인 순차 학습 알고리즘(Online sequential learning algorithm, 이하 OSLA)은 Tikhonov-Phillips 함수를 사용하여 가중값 초기화에 대한 생성 종속성을 해결할 수 있다. OSLA는 그동안의 ELM 알고리즘보다 더 개선된 학습 모델이다. George와 Mutalik(2019)의 연구에서 제공된 OSLA에 대한 MMST(modified minimum slack time) 최적화는, 실행 가능한 최적화 기술을 사용하여 더 많은 개선이 이루어질 수 있음을 보여준다. 이 연구에서는 SLFNN을 사용한 간단한 OSLA를 제시하고 있고, 알고리즘의 수렴 속성은 재귀적 최소 제곱 알고리즘과 유사하다. 예측 성능은 다른 OSLA보다 우수하며 통계적으로 차별된 OSLA 임을 입증하고 있다. 또한 학습 알고리즘의 예측 성능을 향상시키는 다중 모델, 스위칭 및 튜닝 방법론도 제시하고 있다.

#### 4.2.7 GA-KELM

지능형 교통관리 시스템 등과 같이 예측 정확도와 일반화 성능 향상을 위해서는, 과거 데이터를 분석한 후 몇 분 단위의 교통상황을 신속하고 정확하게 예측해야 하는 문제가 무엇보다 중요하다. 그러나 교통 흐름 데이터는 그 불규칙성과 비선형적 특성 그리고 대규모 및 지속적이고 폭발적인 온라인 데이터 발생으로 인하여 효율적인 예측 모형을 개발하기 어렵다. 따라서 Chai 등(2023)은 예측 정확도와 일반화 성능의 향상을 위하여 GA를 이용한 개선된 커널 익스트림 학습 기계(Genetic Algorithm improved Kernel ELM, 이하 GA-KELM)를 제안하였다. 커널 함수를 이용하는 유전자 학습 모형인 GA-KELM은 기존 ELM을 이용한 단기 교통 흐름 예측의 정확도를 상대적으로 향상시키고 있다.

#### 4.2.8 IG-ELM

ELM은 일반적으로 배치 학습 모드를 위한 분류기이다. 반면에 OS-ELM 알고리즘은 데이터를 1:1 또는 청크 단위로 읽는 순차 입력을 처리하도록 진행된다. OS-ELM의 무작위 초기화와 중복되고 관련 없는 기능의 존재로 인해 전체 시스템의 일반화 성능이 평가절하될 수 있다. 이러한 문제를 해결하기 위해 Kale(2022)은 임상 데이터 세트를 활용하여 방사형 기반 또는 기능에 대한 순차 입력에 대한 개선된 유전자 최적화 특징 선택 분류기(improved genetic optimized for ELM classifier, 이하 IG-ELM)를 제안하였다. 제안된 모형은 ELM, IG-ELM, OS-ELM, IG-OSELM 등과 함께 평가하였는데, 여기서 Kale(2022)은 ELM과 대비하였을 때 IG-ELM은 43.25% 학습효율의 향상과 10.94% 향상된 예측 정확도를 제공하는 것으로 나타났다.

## V. 결론 및 논의점

시계열 예측과 분석은 다양한 방법으로 수행될 수 있다. 전통적인 시계열 모형들은 추세 학습과 구현에 대한 기초를 제공하지만, 선형성 가정이라는 제약으로 인하여 그 추세를 완벽하게 모델링하지 못하는 한계가 있다. 결과적으로 신경망의 등장으로 인하여 시계열 자료의 비선형 문제를 해결하고 이를 다루기 위한 다양한 방향을 제시할 수 있게 되었다. ANN은 생성된 데이터의 비선형성을 정확한 모형으로 쉽게 매핑할 수 있으며, 여기서 FNN은 다양한 방식으로 개선된 학습을 할 수 있다.

BPA의 오차 기울기 접근 방식은 RNN이나 LSTM과 같은 신경망 모형으로의 적용 가능성을 확대시켰다. 이러한 모형은 시계열 자료의 유사한 패턴을 구별할 수 있으며, 시간 간격이 다를 수 있는 시계열의 특성을 잘 파악한다. 특히 LSTM은 과거의 추세를 현재 데이터와 연결하여 보다

더 정확한 결과를 생성하며, RNN과 LSTM은 시계열 자료의 비선형 문제를 효과적으로 해결하는 데 기여하고 있다.

RNN을 이용한 신경망 학습은 일반적으로 학습 시간이 오래 걸리고 외생변수 속성을 처리하기 어려운 단점이 있다. 이러한 문제는 ELM을 활용하면 극복할 수 있다. ELM은 가중값 갱신을 통하여 학습을 빠르고 효율적으로 수행할 수 있는 방법으로 알려져 있다. OS-ELM과 그 개선된 확장 기법들은 온라인 시계열 자료 모델링에 있어 정확하고 빠른 접근법을 제시하고 있다. 특히 GA-KELM은 지속적이고 폭발적으로 발생하는 불규칙적이며 비선형적 특성의 대규모 온라인 데이터를 학습시킬 수 있고 효율적인 예측 모델의 개발에 적용할 수 있다. 그러나 초기 가중값의 영향으로 포화 학습 결과를 보일 수도 있다. 초기 가중값 선택에 독립적인 문제 해결 방법을 갖고 있는 OSLA 모형은 이 문제를 해결할 수도 있다. 따라서 OSLA 모형은 비선형 학습의 대안이 될 수 있다.

또 다른 고려해야 할 시계열 자료의 문제는 외생변수 요인을 포함하고 있는 경우이다. 이러한 경우에는 FIS을 이용한 OSF-ELM이 효과적일 수 있다. 이 방법은 외생변수를 시계열 분석에 통합하며, OSLA와 OS-ELM 및 그 확장 연구들을 통하여 시계열 예측에 우수한 결과를 제시하고 있다. 이러한 모델은 MMST나 PSO와 같은 최적화 기술을 활용하여 더욱 개선시킬 수 있다. 또한 최적화 분야의 군집 지능(Swarm Intelligence)은 향후의 주요 연구 분야로서 이를 ELM에 적용하여 예측 성능을 향상시키는데 활용할 수 있을 것이다. 뿐만 아니라 외생변수 입력과 온라인 데이터를 포함하는 확장 연구 또한 앞으로의 연구 과제로 기대되는 분야이다.

## 참 고 문 헌

- [1] Han M, Zhang R, Qiu T, Xu M and Ren W, "Multivariate Chaotic Time Series Prediction Based on Improved Grey Relational Analysis," IEEE Transactions on Systems, Man, and Cybernetics: Systems, pp. 1-11, 2017.
- [2] O'Reilly C, Moessner K and Nati M, "Univariate and Multivariate Time Series Manifold Learning," Knowledge-Based Systems, 133, pp. 1-16, 2017.
- [3] Wilson S, "Data representation for time series data mining: Time domain approaches," Wiley Interdisciplinary Reviews: Computational Statistics, 9(1), p. e1392, 2016.
- [4] Gooijer JD and Hyndman R, "25 years of time series forecasting," International Journal of Forecasting, 22(3), pp. 443-473, 2006.
- [5] Yi BK and Faloutsos C, "Fast Time Sequence Indexing for Arbitrary Lp Norms," Proceedings of the VLDB, Cairo, Egypt, 2000.
- [6] Lin J, Keogh E, Wei L and Lonardi S, "Experiencing SAX: A novel symbolic representation of time series," Data Mining and Knowledge Discovery, 15(2), 2007.
- [7] Fayyad UM and Irani KB, "Multi-Interval Discretization of continuous-valued Attributes for Classification Learning," Proc. 13th Int'l Joint Conference of Artificial Intelligence, pp.1022-1027, 1993.
- [8] George K and Mutalik P, "A Multiple Model Approach to Time-Series Prediction Using an Online Sequential Learning Algorithm," IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(5), pp. 1-15, 2019.
- [9] Rosenblatt F, "The Perceptron: A Probabilistic Model for Information Storage and Organization

- in the Brain,” *Psychological Review*, 65(6), pp. 386-408, 1958.
- [10] Huang G, Zhu Q and Siew C, “Extreme learning machine: A new learning scheme of feedforward neural networks,” *IEEE International Joint Conference on Neural Networks*, 2004.
- [11] Tang J, Deng C and Huang G, “Extreme Learning Machine for Multilayer Perceptron,” *IEEE Transactions on Neural Networks and Learning Systems*, 27(4), 2016.
- [12] Kuo J, Principe J and Vries BD, “Prediction of chaotic time series using recurrent neural networks,” *Neural Networks for Signal Processing II Proceedings of the 1992 IEEE Workshop*, 1992.
- [13] Li Q and Lin RC, “new approach for chaotic time series prediction using recurrent neural network,” *Mathematical Problems in Engineering*, 2016.
- [14] Qinghai L, Lin RH, “A New Approach for Chaotic Time Series Prediction Using Recurrent Neural Network,” *Mathematical Problems in Engineering*, 2016.
- [15] Ho S, Xie M and Goh T, “A comparative study of neural network and Box-Jenkins ARIMA modeling in time series,” *Computers & Industrial Engineering*, 42(2-4), pp. 371-375, 2002.
- [16] Gers F, Eck D and Schmidhuber J, “Applying LSTM to Time Series Predictable Through Time-Window Approaches,” *Perspectives in Neural Computing*, pp. 193-200, 2002.
- [17] Hou X, Wang K, Zhang J and Wei Z, “An Enriched Time-Series Forecasting Framework for Long-Short Portfolio Strategy,” *IEEE Access*, 8, 2020.
- [18] Markowska-Kaczmar U and Kosturek M, “Extreme learning machine versus classical feedforward network,” *Neural Computing & Applications*, 33(22), pp. 15121-15144, 2021.
- [19] Huang G, Wang D and Lan Y, “Extreme learning machines: A survey,” *International Journal of Machine Learning and Cybernetics*, 2(2), pp. 107-122, 2011.
- [20] Ahmad I, Basher M, Iqbal M and Rahim A, “Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection,” *IEEE Access*, 6, 2018.
- [21] Wang Z, Li M, Wang H, Jiang H, Yao Y, Zhang H, Xin J, “Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features,” *IEEE Access*, 7, 2019.
- [22] Smith M, *Neural Networks for Statistical Modeling*, Van Nostrand Reinhold, 1996.
- [23] Bharath, YK, Griffiths’ Variable Learning Rate Online Sequential Learning Algorithm for Feed-Forward Neural Networks, *Automatic Control and Computer Sciences; New York*, 56(2), pp. 160-165, 2022.
- [24] Rajendra V and Motupalli R, “Analysis on the Estimators to the OSELM Model,” *4th International Conference on Advances in Computing, Communication Control and Networking (ICAC 3N)*, 2022.
- [25] Chang HD, Wang XJ, Gu J, Wang W, “A Review of Online Sequential Extreme Learning Machines,” *Journal of Physics: Conference Series*, 1302(3), 2019.
- [26] Sun Z, Au K and Choi T, “A Neuro-Fuzzy Inference System Through Integration of Fuzzy Logic and Extreme Learning Machines,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(5), 2007.
- [27] Zhou X, Liu Z and Zhu C, “Online Regularized and Kernelized Extreme Learning Machines with

- Forgetting Mechanism,” Mathematical Problems in Engineering, 2014.
- [28] Li R, Zhang H, Wu Z, Li R, “Transient Electromagnetic Inversion: An ICDE-Trained Kernel Principal Component OSELM Approach,” IEEE Transactions on Geoscience and Remote Sensing, 60, pp.1-14, 2022.
- [29] George K, Mutalik P, “A Multiple Model Approach to Time-Series Prediction Using an Online Sequential Learning Algorithm,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(5), 2019.
- [30] Scardapane S, Comminiello D, Scarpiniti M and Uncini A, “Online Sequential Extreme Learning Machine With Kernels,” IEEE Transactions on Neural Networks and Learning Systems, 26(9), pp. 2214-2220, 2015.
- [31] Lu J, Huang J and Lu F, “Time Series Prediction Based on Adaptive Weight Online Sequential Extreme Learning Machine,” Applied Sciences, 7(3), 2017.
- [32] Atsawaraungsuk S, Boonphairote W, Somsuk K, Suwannapong C, Khummanee S, “A progressive learning for structural tolerance online sequential extreme learning machine,” TELKOMNIKA Telecommunication Computing Electronics and Control, 21(5), pp. 1039-1050, 2023.
- [33] Matias T, Souza F, Araujo R, Goncalves N and Barreto J, “On-line sequential extreme learning machine based on recursive partial least squares,” Journal of Process Control, 27, pp. 15-21, 2015.
- [34] Rong HJ, Huang GB, Sundararajan N and Saratchandran P, “Online Sequential Fuzzy Extreme Learning Machine for Function Approximation and Classification Problems, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39(4), pp. 1067-1072, 2009.
- [35] Chai W, Zheng Y, Lin, T, Qin J, Zhou T, “GAKELM: Genetic Algorithm Improved Kernel Extreme Learning Machine for Traffic Flow Forecasting,” Mathematics, 11(16), 2023.
- [36] Kale AP, Sonawane S, Wahul RM, Dudhedia MA, “Improved Genetic Optimized Feature Selection for Online Sequential Extreme Learning Machine, Ingenierie des Systemes d’Information,” 27(5), pp. 843-848, 2022.

### 저자 소개



#### 윤여창(YeoChang Yoon)

- 1986년 2월: 성균관대학교 통계학과 (학사)
  - 1989년 8월: 성균관대학교 통계학과 (석사)
  - 1996년 2월: 성균관대학교 통계학과 (박사)
  - 1995년 3월~현재: 우석대학교 정보보안학과 교수
- <관심분야> 인공지능망 모델링 및 학습, 시계열 분석, 빅데이터 분석