

LSTM/GRU 순환신경망을 이용한 시계열 데이터 예측

김 호 현

한국방송통신대학교 대학원

hohkim@koreanair.com

Forecasting Time-series Data Using LSTM/GRU Recurrent Neural Networks

Hohyun Kim

Korea National Open University Graduate School

요 약

주가지수, 환율 등과 같은 시계열자료에 대한 정확한 예측은 많은 이해관계자들의 주요 관심사로서 많은 연구가 진행되어 왔다. 전통적으로 통계적인 ARIMA 기법이 많이 사용되어 왔으나, 인공신경망을 이용한 예측에 관한 연구가 활발하게 이루어져 왔고, 예측 정확도를 더욱 높이기 위하여 ARIMA와 인공신경망을 결합한 Hybrid Model에 관한 연구도 시도되어 왔다. 최근 딥러닝 분야의 급속한 진보는 이미지 처리, 음성인식, 자연어 처리 등의 분야에서 커다란 성과를 가져오고 있다. 본 논문에서는 인공지능의 최근성과를 시계열자료 예측 문제에 적용한다. 순환신경망은 입력데이터의 선후관계를 인식할 수 있고, LSTM/GRU는 긴 문맥의 처리를 가능케 한다. LSTM/GRU 순환신경망을 시계열자료 예측 문제에 적용 시 기존의 Hybrid Model을 능가하는 성능을 얻을 수 있다. 본 논문은 그 방법론을 제시하며, 실험결과는 제안모델이 기존 연구결과를 능가하는 성능을 보임을 입증한다.

1. 서론

시계열데이터는 도처에서 발견된다. 주가, 환율, 유가, 물가, 매출액, 방문객수 등이 좋은 예이다. 기업, 관공서, 금융기관을 포함한 많은 이해관계자들에게 시계열데이터는 매우 중요하며, 그것에 대한 예측은 항상 주요한 관심사이다. 따라서 시계열데이터의 예측에 관한 많은 연구가 있어왔다.

- 전통적 통계기반 ARIMA: 선형관계
- 인공신경망 기반 ANN: 정상성(stationarity)나 데이터에 분포에 대한 가정이 필요 없음. 선형관계, 비선형관계를 학습 가능
- Hybrid Model: 상기 둘을 결합하여 예측력을 높이려는 시도. ARIMA와 ANN을 독립적으로 적용했을 때 보다 높은 성능을 보임.

최근에 와서 인공지능의 진보가 눈부시다. 최근의 자율주행, 음성인식, 이미지 처리 등을 포함한 산업 전 분야에서 일어나고 있는 비약적인 성과의 배경에는 인공지능이 자리하고 있으며,

인공지능의 성과는 진보한 인공신경망, 딥러닝에 힘입은 바 크다.

최근에는 전통적인 FCNN을 넘어 컨볼루션 인공신경망(convolution neural networks), 순환신경망(recurrent neural networks)에 대한 연구가 활발하다. 본 연구에서는 데이터의 선후관계를 인식하여 처리할 수 있는 순환신경망에 집중한다. 시계열데이터는 시간 순서에 따라 발생하는 데이터이므로 선후관계의 문맥을 이해하는 순환신경망은 자연스러운 선택이며, 태생적으로 시계열데이터에 강점을 갖는다고 생각할 수 있다.

2. 관련 연구

2.1 ARIMA

- ARIMA에 대한 일반적 설명
- ARIMA 수식

국내에서는 2014년에 ARIMA 모형을 활용하여 저가항공의 수요를 예측하는 연구가 있었다 [10].

2.2 Neural Network

Fully Connected Neural Networks. 통계적 기법들이 요구하는 데이터에 대한 어떠한 가정도 필요없이, 데이터 자체에만 의존하여 학습을 통해 예측을 할 수 있다는 강점이 있다.

여러 연구가 있었지만 실제 그다지 만족할 만한 성능을 보이지 못하였다.

Chaudhuri et al. 은 Multi layer feed forward 신경망과 NARX(Nonlinear Autoregressive models with exogenous input) 신경망을 이용하여 환율을 예측하는 연구를 진행하였다.[2]

Wang et al. 은 2016년 Elman 순환신경망을 활용하여 Financial Time Series 예측과 원유가격 예측 연구를 진행하였다. [8] [7]

Jalal et al. 의 Elman과 NARX 신경망을 결합한 연구사례도 있다.

국내에서는 2014년에 ARIMA와 신경회로망을 비교 실험하여 상품 수요예측 모형을 개발하는 연구가 있었다 [11].

2.3 ARIMA와 신경망을 혼합한 Hybrid Model

선형관계를 잘 파악할 수 있는 ARIMA와 비선형관계를 학습할 수 있는 인공신경망(Artificial Neural Network)을 결합하여 예측성능을 높이려는 여러 연구가 있어왔다.

1. Zhang's hybrid Model, 2003[9]
2. Khashei & Bijari's hybrid Model, 2010[5]
3. Babbu & Reddy's hybrid Model, 2014[1]

3. 우리의 모델

3.1 Recurrent Neural Networks

3.2 Simple RNN

그림은 torch7 webinar video 33:35 참조

- for modeling sequential data like text, speech, videos
- 3 layers: input(V), recurrent(U) and output(W) layer
- feed the previous state as input to next state
- long sequences suffer from exploding and vanishing gradients

3.3 BPTT: Back-propagation through time

그림은 torch7 webinar video 35:06 참조

- forward-propagate for rho time-steps
- unfold network for rho time-steps
- bak-propagate through unfolded network
- accumulate parameter gradients (sum over time-steps)

3.4 LSTM / GRU

LSTM(Long Short Term Memory)과 GRU(Gated Recurrent Unit)에 대한 설명을 참조하라 [3].

3.5 예측 모델

시계열데이터를 LSTM/GRU RNN에 입력하여, 최종 시계열 데이터의 다음 데이터를 예측한다. 두 가지 방법을 고려해 본다.

1. 데이터를 선처리 없이 그대로 입력 데이터로 사용하는 방법
2. 데이터를 이동평균 필터를 사용하여 Decomposition 하여, 추세데이터와 나머지데이터로 분리한 후, 이들을 입력 데이터로 사용하는 방법

4. 실험

본 논문에서는 다음 세 가지의 데이터셋에 대한 실험을 통해, 지금까지 제안된 여러 모델들과의 성능을 비교 평가한다.

4.1 Deep learning frameworks

현재 일반적으로 많이 사용되고 있는 딥러닝 프레임워크로는 다음과 같은 것들이 있다.

1. Caffe: 버클리대학교
2. Torch: 뉴욕대학교, LuaGIT, Twitter, Facebook, DeepMind [6]
3. Theano: 그래프 기반
4. Tensorflow: 그래프 기반, 구글 [4]

본 저자는 딥러닝의 세부사항을 조금 더 잘 이해할 수 있는 Torch를 기본으로, Tensorflow를 보완적으로 이용하여 실험을 진행하였다.

4.2 Sunspot data 예측

	MAE	REMARKS
-----	-----	-----
ARIMA	13.0337	ARIMA (9, 0, 0)
ANN	12.8918	ANN_ (9x9x1)
Zhang's hybrid	12.7802	
Kashei \& Bijri's hybrid	12.1180	
My Model	11.2100	
Babu \& Reddy		

4.3 환율 예측

4.4 계좌입금액 예측

5. 결론

순환신경망은 입력데이터의 시간관계를 인식할 수 있으므로 전통적인 Fully Connected 인공신경망 보다 시계열데이터에 강점이 있다. LSTM/GRU 셀을 은닉층으로 사용하는 순환 신경망은 Vanishing Gradient 문제를 해결할 수 있으므로 긴 문맥을 문제없이 처리할 수 있다. LSTM/GRU 순환신경망을 이용한 시계열데이터 예측은 ARIMA와 인공신경망을 결합한 Hybrid Model의 성능을 뛰어 넘는다. 향후 FCNN과 RNN의 결합을 포함한 보다 다양한 Topology의 신경망에 대한 연구를 통해 한층 높은 예측력을 갖는 신경망모델을 만들 수 있을 것으로 기대한다.

참고 문헌

- [1] C Narendra Babu and B Eswara Reddy. A moving-average filter based hybrid arima-ann model for forecasting time series data. *Applied Soft Computing*, 23:27–38, 2014.
- [2] Tamal Datta Chaudhuri and Indranil Ghosh. Artificial neural network and time series modeling based approach to forecasting the exchange rate in a multivariate framework. *arXiv preprint arXiv:1607.02093*, 2016.
- [3] Colah. Understanding lstm networks. URL <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- [4] Google. Tensorflow: Open source software library for machine intelligence. URL <https://www.tensorflow.org/>.

- [5] Mehdi Khashei and Mehdi Bijari. An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with applications*, 37(1):479–489, 2010.
- [6] NYU. Torch: A scientific computing framework for luajit. URL <http://torch.ch/>.
- [7] Jie Wang and Jun Wang. Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations. *Energy*, 102:365–374, 2016.
- [8] Jie Wang, Jun Wang, Wen Fang, and Hongli Niu. Financial time series prediction using elman recurrent random neural networks. *Computational Intelligence and Neuroscience*, 2016, 2016.
- [9] G Peter Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50: 159–175, 2003.
- [10] 김영주. 계절 arima 모델을 활용한 저가항공 수요예측 - 국내선 제주권, 내륙권 예측 비교 분석. *관광연구논총*, 26 (1):3–25, 2 2014.
- [11] 이재현. 웨이블렛 신경회로망을 이용한 상품 수요 예측 모형에 관한 연구. *JKIECS*, 9 2014.