

양방향 LSTM 순환신경망 기반 주가예측모델

주일택*, 최승호*

Stock Prediction Model based on Bidirectional LSTM Recurrent Neural Network

Il-Taeck Joo*, Seung-Ho Choi*

요약 본 논문에서는 시계열 데이터인 주가의 변동 패턴을 학습하고, 주가 가격을 예측하기 적합한 주가 예측 딥러닝 모델을 제시하고 평가하였다. 일반신경망에 시계열 개념이 추가되어 은닉계층에 이전 정보를 기억시킬 수 있는 순환신경망이 시계열 데이터인 주가 예측 모델로 적합하다. 순환신경망에서 나타나는 기울기 소멸문제를 해결하며, 장기 의존성을 유지하기 위하여, 순환신경망의 내부에 작은 메모리를 가진 LSTM을 사용한다. 또한, 순환신경망의 시계열 데이터의 직전 패턴 기반으로만 학습하는 경향을 보이는 한계를 해결하기 위하여, 데이터의 흐름의 역방향에 은닉계층이 추가되는 양방향 LSTM 순환신경망을 이용하여 주가예측 모델을 구현하였다. 실험에서는 제시된 주가 예측 모델에 텐서플로우를 이용하여 주가와 거래량을 입력값으로 학습을 하였다. 주가예측의 성능을 평가하기 위해서, 실제 주가와 예측된 주가 간의 평균 제곱근 오차를 구하였다. 실험결과로는 단방향 LSTM 순환신경망보다, 양방향 LSTM 순환신경망을 이용한 주가예측 모델이 더 작은 오차가 발생하여 주가 예측 정확성이 향상되었다.

Abstract In this paper, we proposed and evaluated the time series deep learning prediction model for learning fluctuation pattern of stock price. Recurrent neural networks, which can store previous information in the hidden layer, are suitable for the stock price prediction model, which is time series data. In order to maintain the long-term dependency by solving the gradient vanish problem in the recurrent neural network, we use LSTM with small memory inside the recurrent neural network. Furthermore, we proposed the stock price prediction model using bidirectional LSTM recurrent neural network in which the hidden layer is added in the reverse direction of the data flow for solving the limitation of the tendency of learning only based on the immediately preceding pattern of the recurrent neural network. In this experiment, we used the Tensorflow to learn the proposed stock price prediction model with stock price and trading volume input. In order to evaluate the performance of the stock price prediction, the mean square root error between the real stock price and the predicted stock price was obtained. As a result, the stock price prediction model using bidirectional LSTM recurrent neural network has improved prediction accuracy compared with unidirectional LSTM recurrent neural network.

Key Words : Bidirectional, Deep learning, LSTM, Long Short-Term Memory, Prediction, Recurrent Neural Network(RNN), Stock Price

1. 서론

과거의 주식 가격과 거래량 등의 관계를 시계열 분

석하여 미래의 주식 가격을 예측한다. 주식 가격의 데이터는 시간에 따른 변동성을 가지므로, 과거의 데이터를 ARCH, GARCH 등의 시계열 분석 모델로 주가를

*Department of Computer Science, Dongshin University
 Received April 09, 2018

Revised April 11, 2018

Accepted April 19, 2018

예측을 하였다[1][2]. 그러나 어느 특정 시점의 측정치와 그 이전의 측정치 관계가 비선형적이고 경향에 종속적이기 때문에 기존에 통계적인 분석 방법으로 접근하기보다는 신경망으로 접근하는 것을 제안하였다[3]. 다차원 공간에 복잡한 비선형 분류가 가능한 기계학습인 순환신경망의 딥러닝 알고리즘은 컴퓨팅 기술의 빠른 발전에 따라 주가 예측에서 사용되고 있다[4][5]. 순환신경망은 내부에 정보가 지속되도록 순환구조를 가진 시계열 개념이 추가된 것으로, 은닉계층에 과거의 데이터를 기억하여 학습을 진행하므로 데이터의 시간적 순서 관계가 중요한 시계열 데이터에 적합한 모델이다. 그러나 순환신경망은 절차가 길어지면 어느 단계부터 학습이 더 이루어지지 않게 되는 기울기 소실이 발생하면서, 과거의 학습 결과가 사라지는 장기 의존성 문제가 있다. 기울기가 사라지는 문제를 극복하기 위해서, 제안된 장기 의존성을 학습할 수 있는 순환신경망 구조의 변형인 LSTM (Long Short-Term Memory)은 1997년에 제안되었다[6].

순환신경망의 성능을 개선하기 위해 다양한 변형모델이 제안되고 있다. 시계열 데이터의 역방향에 은닉계층을 추가하여 성능이 향상되는 양방향 순환신경망이 제안되었다[8]. 본 논문에서는 양방향 LSTM 순환신경망을 이용하여 주가를 예측하는 모델을 제시하고 구현하여 성능평가를 하였다.

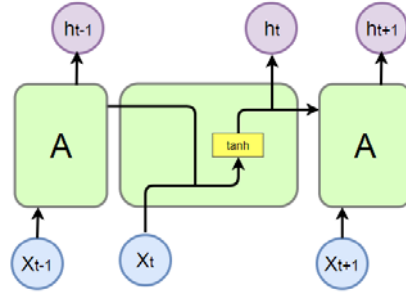
본 논문의 구성은 다음과 같다. 제2장에서는 주가예측을 위한 딥러닝 모델에 대해 설명한다. 제3장에서는 본 논문에서 제시한 주가예측 딥러닝 모델에 대한 실험 및 평가 결과 내용에 대해 설명한다. 마지막으로, 제4장에서는 결론 및 향후 연구를 제시한다.

2. 주가예측 딥러닝 모델

2.1 순환신경망

앞에서 일어난 과거의 주가가 미래의 주가에 어떻게 영향을 미치는지를 예측하는 딥러닝 모델 중에서 순환신경망(Recurrent Neural Network)은 일반신경망에서 시계열 개념이 추가된 것으로 은닉계층에 이전 정보를 기억시킬 수 있는 장점이 있다. 순환신경망은 내부에 정보가 지속되도록 순환구조를 가진 시계열 개념이

추가된 것으로 은닉계층에 과거의 데이터를 기억하여 학습을 진행하므로, 데이터 시간적 순서 관계가 중요한 시계열 데이터에 적합한 모델이다.



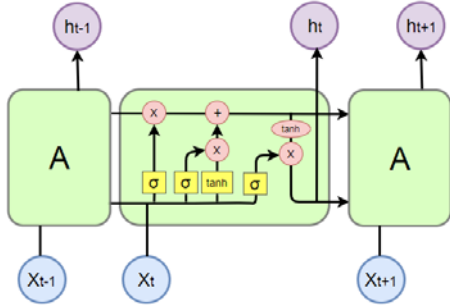
1.
Fig. 1. Recurrent Neural Network(RNN)

그림 1에서는 순환신경망의 순환구조를 펼쳐진 형태로 보여주고 있다. 순환신경망은 입력데이터와 이전 절차 단계 결과를 함께 입력받아 처리하는 절차를 반복한다. 따라서 과거의 데이터가 다음의 데이터에 어떤 영향을 주는지를 학습할 수 있고, 이러한 장점으로 인해 순환신경망은 입력 데이터가 시간 순서로 주어지는 주가의 시계열 데이터 예측 등에 사용한다. 순환신경망은 입력을 비선형 관계로 변환하는 하이퍼볼릭 탄젠트(Tanh)의 활성화 함수를 사용하여 출력을 내며, 다음 절차 단계의 입력으로 전해진다. 출력 값은 -1과 1사이의 작은 값을 갖기 때문에, 출력 값들이 곱해지는 단계가 반복되면, 숫자가 점점 작아진다. 이러한 절차가 길어지면 어느 단계부터는 학습이 더 이루어지지 않게 되는 기울기 소실이 발생하면서, 과거의 학습 결과가 사라지는 장기 의존성 문제가 있다.

2.2 LSTM 순환신경망

LSTM(Long Short-Term Memory)은 기울기가 사라지는 문제를 극복하기 위해서 제안된 장기 의존성을 학습할 수 있는 특별한 종류의 순환신경망 구조이다. 사라지는 기울기의 문제를 극복하기 위해서, LSTM 구조는 순환신경망의 일반적인 신경들을 내부에 작은 메모리를 가진 LSTM 셀(Cell)로 교체한다. 이 LSTM 셀들은 일반적인 순환신경망과 같이 연결되어 있어서,

셀 내부의 상태가 여러 단계에 걸친 오류를 기억하는 것을 도와준다.



2. LSTM(Long Short-Term Memory)
Fig. 2. LSTM Recurrent Neural Network

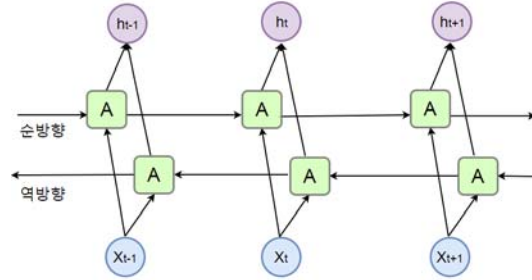
그림 2에서의 LSTM 셀은 상태를 유지하기 위해서, 다음과 같은 게이트(Gate)들로 이루어져 있다. 입력 게이트와 망각 게이트 및 출력 게이트 등을 이용하여, 출력 값을 조절한다. 입력 게이트는 입력 값을 얼마나 받아들일지를 결정하고, 망각 게이트는 이전의 셀 내부의 상태를 얼마나 잊어버릴지를 결정하며, 출력 게이트는 얼마나 출력할지를 결정한다. 이처럼 셀 상태 값을 얼마나 잊어버리고, 새로운 입력 값을 얼마나 받아들일지 결정하여 더하는 구조이다. 이 과정이 반복되어도 기울기가 소멸하여 학습할 수 없어지는 일이 발생하지 않는다.

LSTM은 표준 순환신경망과 같은 방법으로 은닉 변수를 거쳐 최종 출력 값을 계산하지만, 은닉 계층의 변수 계산 과정에서 게이트들을 적절하게 이용하여 정보의 흐름을 조절한다. 이러한 결과로 LSTM 셀을 사용한 순환신경망은 주식 시계열데이터와 같은 긴 절차 순서의 데이터라도 기울기 소실 문제없이 처리한다.

2.3 양방향 순환신경망

일반적으로 순환신경망은 일반신경망 학습과 달리 시계열 개념이 추가된 것으로 은닉 계층에 과거의 데이터의 정보를 기억할 수 있다. 그러나 입력 순서를 시간 순서대로 처리하기 때문에 결과물은 주로 직전 패턴을 기반으로 하는 경향을 보이는 한계가 있다[7]. 이러한

한계를 해결하기 위해서, 양방향 순환신경망은 순방향과 역방향의 두 개의 분리된 순환신경망을 통해 학습을 시킨다. 이러한 방법으로, 학습시킬 시계열 데이터의 역방향에 추가로 은닉계층을 추가시킨다면 성능이 향상한다[8].



3.
Fig. 3. Bidirectional Recurrent Neural Network

일반적인 순환신경망은 입력과 출력 계층에 순방향의 은닉 계층만 존재하지만, 그림 3에서의 양방향 순환신경망은 입력과 출력 계층에 순방향 은닉 계층과 추가로 역방향 은닉 계층이 연결된 구조를 보여주고 있다.

2.4 제안 방법과 관련한 내용

본 논문에서 주가 예측을 위한 딥러닝 모델은 양방향 순환신경망에 LSTM을 사용하였다. 네트워크 전체적인 구조는 주식의 일별 시가, 고가, 저가, 종가, 거래량을 입력으로 입력계층은 5개로 구성되어 있으며 20개의 LSTM 셀로 구성된 은닉계층과 주가 예측을 위한 1개 출력계층으로 이루어진다.

딥러닝의 최적화기(Optimizer)로 아담(Adam) 최적화기를 이용하여, 과거 및 현재의 주가와 거래량을 이용하여 미래의 주가를 학습한다.

3. 실험 및 성능 평가

본 장에서는 제시된 딥러닝 모델을 구현하고 실험한 환경 및 모델의 학습, 테스트, 검증 결과에 관해 설명한다.

3.1 실험 환경

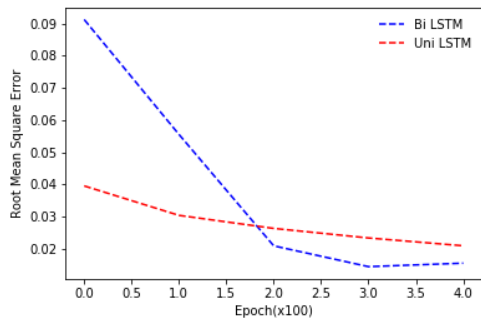
본 논문에서는 주가 데이터는 구글(Google)의 주가

를 선택하였으며, 분석 데이터는 시가, 고가, 저가, 종가, 거래량 등의 기본 데이터를 사용하였다. 전체 주가 데이터 3428개의 70%를 학습용 데이터, 나머지 30%를 테스트용 데이터로 이용하여 실험을 진행하였다. 주가 예측의 오차율을 확인하기 위해서 다음 식(1)과 같이 실제 주가와 예측된 주가 간의 평균 제곱근 오차(Root Mean Square Error)를 사용하였다.

$$RMSE = \sqrt{\frac{1}{n} \left(\sum_{t=1}^n (\text{real} - \text{predict}_t)^2 \right)} \quad \text{식(1)}$$

3.2 성능평가

본 실험은 Tensorflow v.1.7에서 Learning rate는 0.01로 설정하고 500번의 반복 학습을 수행하였다. 실험을 통해서 제시된 딥러닝 모델에 대해 오차 감소율과 예측률을 각각 측정하였다.



4. LSTM (RMSE)

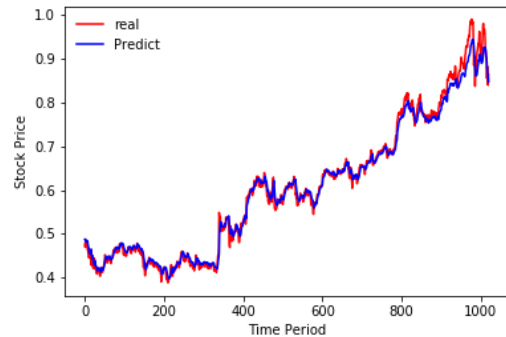
Fig. 4. Comparison of RMSE results between unidirectional and bidirectional LSTM RNN

그림 4는 단방향(Unidirectional) LSTM 순환신경망과 양방향(Bidirectional) LSTM 순환신경망 모델을 학습이 진행되면서 발생하는 평균 제곱근 오차(RMSE)를 출력해 그래프로 나타냈다. 여기에서 학습이 진행될수록 양방향 LSTM 순환신경망 모델은 단방향 LSTM 순환신경망 모델 보다 평균 제곱근 오차가 더 낮아지는 결과를 보여주고 있다.

1. (RMSE)
Table 1. Results of RMSE and Prediction rate

	LSTM	LSTM
RMSE	0.0209	0.015
	83%	84.9%

양방향 LSTM 순환신경망의 최소 RMSE가 0.015이며, 단방향 LSTM 순환신경망보다 작은 오차 값을 보였고 주가 예측률도 단방향 LSTM 모델보다 향상된 성능을 보이고 있다.



5. 가 가

Fig. 5. Graph of real stock price and predicted stock price

그림 5는 양방향 LSTM 순환신경망 모델을 적용하여 예측한 주가와 실제 주가를 비교 표시한 그래프이다. 주가 등락 폭이 작으면 예측 오차가 크게 차이가 없는 결과를 보여주고 있다.

4. 결론 및 향후 연구

본 논문에서는 양방향 LSTM 순환신경망을 이용하여 주가 예측 모델을 구현하여 실험을 통해 제안하는 방법의 성능을 측정하였다. 그 결과는 단방향 LSTM 순환신경망을 사용하였을 때보다 양방향 LSTM 순환신경망 모델이 향상된 성능을 갖는 학습 모델로 도출되었다. 현재는 매우 간단한 구조로 구성되어 있고, 단순한 주가와 거래량 데이터 특징만을 이용해 실험을 진행하였다. 앞으로 양방향 LSTM 순환신경망의 다양한 최적화 연구 과정을 통해 더욱 성능이 향상된 학습모델로

발전할 가능성이 높아 질 것으로 생각한다.

REFERENCES

[1] Robert Engle, "GARCH 101: The use of ARCH/GARCH Models in Applied Econometrics," Journal of Economic Perspectives, vol. 15, no. 4, pp. 157-168, 2001.

[2] V. K. Menon, N. C. Vasireddy, S.A.Jami, V. T. Pedamallu, V. Sureshkumar, K. Soman, "Bulk price forecasting using spark over NSE data set," International Conference on Data Mining and Big Data. Springer, pp. 137-146, 2016.

[3] K. Chakraborty, K. Mehrotra, C.K. Mohan, Sanjay Ranka, "Forecasting the Behavior of Multivariate Time Series using Neural Network," Neural Networks vol. 5, pp. 961-970, 1992,

[4] J. Roman, A. Jameel, "Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns" System Sciences, Proceedings of the Twenty-Ninth Hawaii International Conference on, vol. 2, pp. 454-460, 1996.

[5] E. W. Saad, D. V. Prokhorov, D. C. Wunsch, "Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks," IEEE Transactions on Neural Networks, vol. 9, no. 6, pp. 1456- 1470, 1998.

[6] S. Hochreiter, J. Schmidhuber, J., "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[7] T. Robinson. "An application of recurrent nets to phone probability estimation," IEEE Transactions on Neural Networks, vol. 5 no. 2, pp. 298-305, 1994.

[8] M. Schuster, K. K. Paliwal, "Bidirectional recurrent Neural Networks," IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673-2681, 1997.

(Il-Taeck Joo) []



• 1997 2 :
()
• 1999 2 :
()
• 2008 2 ~ :
()

< > ,

(Seung-Ho Choi) []



• 1992 2 :
()
• 1992 3 ~ :

< >