

Mamba, Transformer, LSTM을 이용한 주식 예측 모델 분석

서세일, 백상렬, 이수안

세명대학교 컴퓨터학부

tpdlf9801@gmail.com, qkqhtkdduf@naver.com, suanlee@semyung.ac.kr

Stock Prediction Model Analysis using Mamba, Transformer and LSTM

Seil Seo, Sangyeol Baek, Suan Lee

School of Computer Science, Semyung University

요 약

최근 인공지능 기술의 발전으로 학습을 통해 복잡한 패턴이나 흐름을 이해하고 예측하는 능력이 크게 향상되고 있다. 주가 데이터는 복잡한 패턴이나 흐름을 지닌 대표적인 데이터이며 주가에 영향을 미치는 변수가 많아 이를 정확하게 예측하는 것은 많은 어려움이 따른다. 본 논문에서는 이러한 문제를 해결하기 위해 Mamba, Transformer, LSTM을 이용하여 주가 데이터 예측 연구를 수행하였다. 세 가지 모델에 주가 데이터를 적용하여 학습하고 예측을 진행하여 성능을 평가하고, 각 모델의 장단점을 분석하였다. 실험 결과는 각 모델의 예측 정확도와 안정성을 비교하여 주가 예측에 가장 적합한 모델에 대한 통찰을 제공한다. 이를 통해 향후 주식 시장 예측 모델 발전에 기여할 수 있을 것으로 기대된다.

1. 서 론

주식 투자는 개인 자산의 수익을 극대화하는 방안으로 좋은 수단이 될 수 있다. 코로나 이후로 투자에 관한 관심이 점차 증가하면서 주식 투자에 관한 관심이 증가하고 있다. 그러나 주식 시장은 불안정하고 불확실한 요소가 많아 주가 예측의 중요성이 커지고 있다. 최근 인공지능 기술이 빠르게 발전하면서 주식 가격 예측 연구가 활발히 진행되고 있으며 모델의 예측 정확도와 최적화를 위한 고도화 작업도 활발히 연구되고 있다.

본 논문은 주식 예측을 위해 학습과 테스트에 쓰일 데이터를 야후 파이낸스를 통해 변동이 심한 미국 주식 시장의 IT 거대 기업인 애플, 아마존, 구글, 메타, 마이크로소프트, 엔비디아, 테슬라로 전체 시점의 주가 데이터를 수집하였다. 시계열 예측의 성능과 특징을 비교하기 위해 딥러닝 모델인 Mamba, Transformer, LSTM(Long Short-Term Memory)을 사용하여 실험을 진행하였고 세 가지 모델 중 Mamba 모델이 7개 기업 주식 데이터에서 학습과 예측을 진행한 성능지표에서 가장 좋은 성능을 확인할 수 있었다.

2. 실험 모델

2.1 Mamba

Mamba[1]는 선택적 상태 공간 모델(Selective State Space Models, SSMs)을 활용한 새로운 신경망 모델이다. 긴 데이터 시퀀스를 효율적으로 모형화하는데 뛰어난 성능을 보인다. Mamba는 입력에 따라 특정 정보에 집중하거나 무시할 수 있는 선택적 집중 메커니즘을 통해 관련 없는 정보를 걸러내고 관련 있는 데이터를 유지할 수 있

다. 또한 합성곱 대신 재귀적으로 계산하는 하드웨어 인식 알고리즘을 사용하여 GPU 메모리 효율성을 극대화하고 계산 속도를 높였다.

기존 Transformer의 계산 복잡도가 시퀀스 길이에 따라 제곱으로 증가하는 문제를 해결하여 시퀀스 길이에 대해 선형적으로 스케일링할 수 있다. 긴 시퀀스에 대해 선형적으로 스케일링할 수 있고, Transformer보다 빠른 추론 속도와 더 작은 메모리를 요구하므로 대규모 모델 학습과 추론에 필요한 컴퓨팅 자원을 절약할 수 있는 것이 특징이며 단순화된 아키텍처로 구현과 확장이 쉬워 다양한 분야에 맞춰 변경하여 사용하기 쉽다는 특징이 있다.

2.2 MambaStock

MambaStock[2]은 Mamba 프레임워크를 활용하여 과거 주식 시장 데이터를 기반으로 미래 주가 변동률을 예측한다. 시작가, 고가, 저가, 거래량 등 다양한 지표를 사용하며 과거 데이터를 처리하여 시간적 종속성을 파악하고 관련 정보를 추출한다. Mamba는 이때 입력 데이터 내의 패턴과 관계를 효과적으로 분석하여 각 날짜에 대한 예측 주가 변동의 차원을 1차원으로 축소한다. 변동률은 $(-1, 1)$ 구간 내에 속하도록 쌍곡선 탄젠트 활성화 함수를 적용한다. 실제 훈련은 예측된 주가 변동률과 실제 주가 변동률의 평균 제곱 오차(Mean Squared Error, MSE)를 사용하여 이를 최소화하여 미래 주가 변동 예측의 정확도를 높이는 것을 목적으로 한다. 이는 높은 예측 정확도와 효

율성을 가지고 주식 시장의 복잡한 패턴을 효과적으로 학습하고 활용할 수 있게 한다.

2.2 Transformer

Transformer[3]는 시계열 데이터나 시퀀스 데이터의 예측에 뛰어난 성능을 발휘하며 입력된 시퀀스의 다음 요소를 예측하는 방식으로 작동한다. 입력 시퀀스의 각 요소에 대한 정보를 파악하고, 다음 요소 예측에 필요한 정보를 추론한다. 이때 어텐션(Attention) 메커니즘과 셀프 어텐션(Self-Attention)을 사용하여 시퀀스의 흐름을 이해하고 중요한 부분에 집중한다.

어텐션 메커니즘은 입력 시퀀스의 모든 위치에 대해 관련성을 계산하여 예측에 중요한 정보를 강조할 수 있게 한다. 셀프 어텐션은 각 요소가 시퀀스 내의 다른 모든 단어와의 관계를 학습하여 문맥을 이해하고 중요한 패턴을 포착한다. 이 과정에서 각 요소의 가중치가 계산되고 이를 기반으로 중요한 정보를 선택적으로 반영된다. 또한 병렬 처리가 가능하여 RNN과 달리 긴 시퀀스도 효율적으로 처리할 수 있다. 학습된 모델은 새로운 입력 시퀀스를 기반으로 다음 시퀀스를 예측할 수 있다. 이러한 특성은 모델이 복잡한 패턴과 장기적인 의존성을 효과적으로 학습할 수 있게 하며 높은 예측 성능을 제공한다.

2.3 LSTM

LSTM[4]은 RNN(Recurrent Neural Network)과는 다르게 긴 시퀀스의 데이터를 처리할 수 있으며 이전 정보를 오랫동안 기억할 수 있어 장기적인 의존성을 효과적으로 처리할 수 있다. LSTM은 입력 게이트, 출력 게이트, 망각 게이트의 세 가지 게이트로 구성되어 있으며, 입력 게이트는 새로운 정보가 셀 상태에 얼마나 영향을 미칠지를 결정하고 출력 게이트는 셀 상태가 다음 상태로 얼마나 전달될지를 조절한다. 망각 게이트는 셀 상태에서 얼마나 많은 정보를 잊을지를 결정하여 불필요한 정보를 걸러내고 중요한 정보만을 유지한다. 이 세 게이트를 사용하여 LSTM이 긴 시퀀스 데이터에서 중요한 패턴을 학습하고 필요한 정보만을 기억하여 다음 처리에 전달할 수 있도록 한다. LSTM의 게이트 메커니즘은 모델이 기억해야 할 정보와 잊어야 할 정보를 스스로 결정할 수 있게 하여, 복잡한 시퀀스 데이터를 효과적으로 처리한다.

3. 실험

3.1 실험 과정

본 논문에서는 주가 예측을 하기 위해 야후 파이낸스 라이브러리를 통해 2012년 1월 1일부터 2024년 4월 28일까지의 기간의 기업들(애플, 아마존, 구글, 메타, 마이크로소프트, 엔비디아, 테슬라)의 주식 데이터 수집했다. 총 3,200개의 데이터 중 처음부터 2,800개의 데이터를 학습 데이터로 사용하였으며 나머지 400개의 데이터를 테스트 데이터로 사용하였다. 이렇게 구성한다면 2,800개 시점이

있는 데이터를 학습하여 전체적인 패턴을 이해하고 그 후 400개의 예측하여 나온 결과와 실제 테스트 데이터와 비교하여 성능 및 정확도가 나온다.

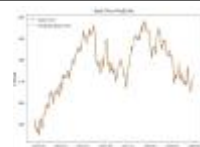
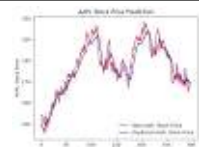

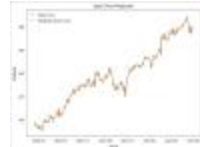


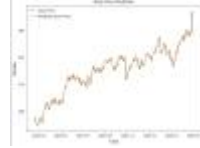
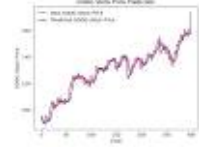




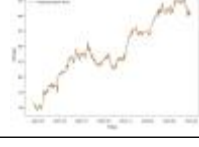
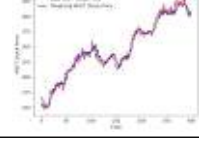

Mamba, Transformer, LSTM을 이용하여 학습과 테스트를 진행하여 결과를 확인하였다.

3.2 실험 결과

MambaStock, Transformer, LSTM 모델의 성능을 비교하고 평가하기 위해 주식 데이터를 사용하여 예측을 진행하였다. 예측된 결과는 표 1과 같이 전체적으로 예측값과 결과값이 유사한 것을 확인하였으며 세 모델의 차이점을 알 수 있었다. MambaStock의 경우 실제 데이터와 예측 데이터가 거의 유사하였고 Transformer의 경우 급격하게 변화될 때 예측이 느리게 반영되는 것을 확인할 수 있었다. LSTM은 전체적으로 추세만 예측이 되는 것을 확인하였고 엔비디아와 같이 심하게 변화할 때 예측하지 못하는 것을 확인하였다.

정확한 수치로 확인하기 위해 표 2, 표 3, 표 4와 같이 MambaStock, Transformer, LSTM 모델의 성능지표를 나열하여 비교분석을 진행하였다. 시각화된 그래프와 성능지표를 종합적으로 분석한 결과 MambaStock이 다른 Transformer와 LSTM에 비해 가장 우수한 성능을 보이는 것으로 나타났다. 특히 예측 정확도와 안정성 측면에서 좋은 성능을 볼 수 있었다.

표 1 주식 예측 결과 그래프

	MambaStock	Transformer	LSTM
애플			
아마존			
구글			
메타			
MS			

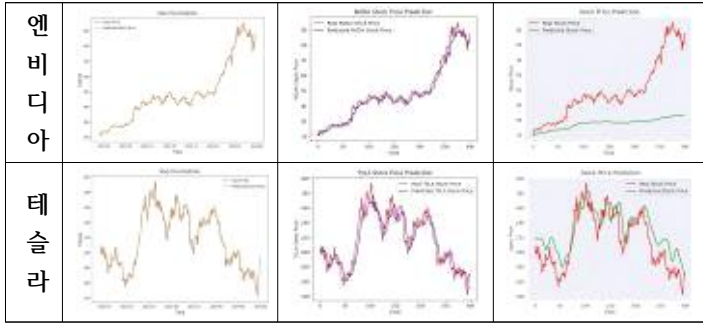


표 3 MambaStock 모델의 기업별 주식 예측 성능지표

기업명	MSE	RMSE	MAE	R ²
애플	4.9768	2.2309	1.7029	0.9662
아마존	6.0913	2.4681	1.8720	0.9909
구글	5.8671	2.4222	1.6813	0.9800
메타	69.5699	8.33409	5.1137	0.9924
MS	22.6637	4.7606	3.7082	0.9903
엔비디아	259.528	16.1099	10.3656	0.9931
테슬라	48.1903	6.9419	5.1208	<u>0.9640</u>
평균	59.5553	6.1811	4.2235	0.9824

표 4 Transformer 모델의 기업별 주식 예측 성능지표

기업명	MSE	RMSE	MAE	R ²
애플	15.1058	3.8866	3.1674	0.8961
아마존	22.3766	4.7304	3.9218	0.9664
구글	14.3749	3.7914	2.8609	0.9504
메타	187.0470	13.6765	9.8172	0.9795
MS	47.5401	6.8949	5.5010	0.9790
엔비디아	8.2168	2.8665	2.1195	0.9780
테슬라	159.6863	12.6367	9.9343	<u>0.8809</u>
평균	64.9068	6.9261	5.3317	0.9472

표 5 LSTM 모델의 기업별 주식 예측 성능지표

기업명	MSE	RMSE	MAE	R ²
애플	64.0906	8.0057	6.8948	0.5590
아마존	52.0217	7.2126	5.9077	0.9218

구글	38.5424	6.2083	4.8213	0.8669
메타	281.2064	16.7692	12.3311	0.9560
MS	105.8723	10.2894	8.5949	0.9533
엔비디아	731.6315	27.0487	21.5594	<u>-0.9624</u>
테슬라	470.5576	21.6923	18.1946	0.6492
평균	249.1318	13.8895	11.1863	0.5634

4. 결론 및 향후 연구

본 논문에서 MambaStock, Transformer, LSTM 모델을 각각 구성하여 같은 주식 데이터를 사용하여 학습과 테스트를 통해 모델별 차이를 알 수 있었다. Mamba 모델의 그래프를 보면 모든 기업의 실제 데이터와 예측값이 거의 일치하였으며 Transformer 모델의 경우 갑자기 급등하거나 하락한 주가를 예측하는데 예측 결과가 느리게 반영되어 예측되었다. LSTM은 대체로 추세는 확인할 수 있었으나 전체적으로 다른 모델에 비해 성능이 떨어지며 엔비디아와 같은 변화가 매우 급격한 경우에는 예측하지 못하였다. 실험 결과는 Mamba 기반의 MambaStock이 그래프와 성능지표를 비교하였을 때 가장 주식 예측 성능이 좋은 것을 알 수 있었다.

본 실험을 통해 MambaStock과 같이 주식 예측을 목적으로 만든 모델의 경우 일반 시계열 예측에 사용되는 모델보다 좋은 성능을 가진 것을 확인할 수 있었다. 규모가 더 크거나 성능이 더 좋은 모델을 사용한다면 정확히 예측하여 더 좋은 성능 결과를 얻을 수 있다고 생각하며 주식 예측에 특화된 모델을 서로 비교하는 것도 의미 있는 연구가 될 것이다. 향후 연구에서는 대규모 언어모델을 사용하여 정확도를 높이는 연구를 진행할 예정이다.

참고문헌

- [1] Gu, Albert, and Tri Dao. "Mamba: Linear-time sequence modeling with selective state spaces." arXiv preprint arXiv:2312.00752 (2023).
- [2] Shi, Zhuangwei. "MambaStock: Selective state space model for stock prediction." arXiv preprint arXiv:2402.18959 (2024).
- [3] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).
- [4] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
- [5] Wu, Haixu, et al. "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting." Advances in neural information processing systems 34 (2021): 22419-22430.